

```
PI Slangenaupt S, Guseella JF;
XX
DR WPI; 2002-674806/72.
XX
PT New IKKAP genes with mutations, useful for identifying a subject with
PT Familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12399 C; 14128 G; 21681 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATCG 20
DB 34222 AGTCGCAACAGTACATCG 34203
XX
RESULT 3
AB080566/c
ID AB080566 standard; DNA; 66479 BP.
XX
AC AB080566;
XX
DT 08-NOV-2002 (first entry)
XX
DE Mutant human IKKAP gene #1.
XX
KW Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW FDI; mutation; gene; chromosome 9q31; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mutation replace(34201,T)
FT /*tag= a
XX
WO200259381-A2.
XX
PD 01-AUG-2002.
XX
PF 07-JAN-2002; 2002MO-US000473.
XX
PR 06-JAN-2001; 2001US-0260080P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
```

```
PI Slangenaupt S, Guseella JF;
XX
DR WPI; 2002-674806/72.
XX
PT New IKKAP genes with mutations, useful for identifying a subject with
PT Familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12399 C; 14129 G; 21680 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATCG 20
DB 34222 AGTCGCAACAGTACATCG 34203
XX
RESULT 4
AB080568/c
ID AB080568 standard; DNA; 66479 BP.
XX
AC AB080568;
XX
DT 08-NOV-2002 (first entry)
XX
DE Mutant human IKKAP gene #3.
XX
KW Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW FDI; PD2; mutation; gene; chromosome 9q31; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mutation replace(33714,G)
FT /*tag= a
FT mutation replace(34201,T)
FT /*tag= b
XX
WO200259381-A2.
XX
PD 01-AUG-2002.
XX
PF 07-JAN-2002; 2002MO-US000473.
XX
PR 06-JAN-2001; 2001US-0260080P.
XX
```

PA (GEHO) GEN HOSPITAL CORP.
XX Staugenhaupt S, Gussella JF;
PI WPI; 2002-674806/72.
XX
PT New IKBAP genes with mutations, useful for identifying a subject with
PT familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKBAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKBAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKBAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKBAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKBAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKBAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12400 C; 14128 G; 21680 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGTCGCAACAGTACATGG 20
Db 34222 AGTCGCAACAGTACATGG 34203
XX
RESULT 5
AB080565/c
ID AB080565 standard; DNA; 66479 BP.
XX
AC AB080565;
XX
DT 08-NOV-2002 (first entry)
XX
XX Human IKBAP wild-type gene.
XX
DE Human; IKBAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW gene; chromosome 9q31; ds.
XX
XX Homo sapiens.
XX OS
XX PN WC020259381-A2.
XX PD
XX 01-AUG-2002.
XX
XX 07-JAN-2002; 2002W0-US000473.
XX
XX 06-JAN-2001; 2001US-0260080P.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX Staugenhaupt S, Gussella JF;
XX
XX WPI; 2002-674806/72.
XX
XX

XX
XX New IKBAP genes with mutations, useful for identifying a subject with
XX PT familial dysautonomia (FD), or for rapid carrier screening in the
XX PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
XX PT prenatal diagnosis.
XX
XX PS Claim 1; Fig 6; 109pp; English.
XX
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKBAP gene (the present
CC sequence) are associated with FD. The mutation associated with the major
CC haplotype of FD, FDI mutation, is a base pair (bp) mutation, where the
CC thymine nucleotide located at bp 6 of intron 20 in the IKBAP gene is
CC replaced with a cytosine. This results in skipping of exon 20 in the mRNA
CC from FD patients, although they continue to express varying levels of
CC wild-type message in a tissue-specific manner. The mutation associated
CC with the minor haplotype, FD2 mutation, is a bp mutation, where the
CC guanine nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a
CC cytosine. This bp mutation causes an arginine to proline missense
CC mutation (R696P) in the IKBAP protein, which is predicted to disrupt a
CC potential phosphorylation site. The IKBAP nucleic acid sequences are
CC useful for identifying a subject with FD and for rapid carrier screening.
CC The IKBAP gene contains 37 exons and maps to chromosome 9q31
XX
SQ Sequence 66479 BP; 18271 A; 12398 C; 14128 G; 21681 T; 0 U; 0 Other;
XX
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGTCGCAACAGTACATGG 20
Db 34222 AGTCGCAACAGTACATGG 34203
XX
RESULT 6
ACF67367.52/c
Continuation (53 of 57) of ACF67367 from base 5200001 (Photorehabus luminescens nucleotic
WP Sequence Split Into 57 Fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000
WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000
WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP

WP ACF67367_31 3100001 3210000
 WP ACF67367_32 3200001 3310000
 WP ACF67367_33 3300001 3410000
 WP ACF67367_34 3400001 3510000
 WP ACF67367_35 3500001 3610000
 WP ACF67367_36 3600001 3710000
 WP ACF67367_37 3700001 3810000
 WP ACF67367_38 3800001 3910000
 WP ACF67367_39 3900001 4010000
 WP ACF67367_40 4000001 4110000
 WP ACF67367_41 4100001 4210000
 WP ACF67367_42 4200001 4310000
 WP ACF67367_43 4300001 4410000
 WP ACF67367_44 4400001 4510000
 WP ACF67367_45 4500001 4610000
 WP ACF67367_46 4600001 4710000
 WP ACF67367_47 4700001 4810000
 WP ACF67367_48 4800001 4910000
 WP ACF67367_49 4900001 5010000
 WP ACF67367_50 5000001 5110000
 WP ACF67367_51 5100001 5210000
 WP ACF67367_52 5200001 5310000
 WP ACF67367_53 5300001 5410000
 WP ACF67367_54 5400001 5510000
 WP ACF67367_55 5500001 5610000
 WP ACF67367_56 5600001 5648894

Query Match Best Local Similarity 84.0%; Score 16.8; DB 10; Length 110000;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTGCACAAACAGTACATGG 20

Db 43881 AGACGCAAAACAGTACAAAG 43862

RESULT 7

Continuation (4 of 7) of ACF65387 from base 300001 (Photographus luminescens nucleotide
 WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP	Fragment Name	Begin	End
WP ACF65387_1	1	110000	110000
WP ACF65387_2	2	210000	210000
WP ACF65387_3	3	300001	310000
WP ACF65387_4	4	400001	510000
WP ACF65387_5	5	500001	610000
WP ACF65387_6	6	600001	696798

Query Match Best Local Similarity 84.0%; Score 16.8; DB 10; Length 110000;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTGCACAAACAGTACATGG 20

Db 83395 AGACGCAAAACAGTACAAAG 83376

RESULT 8

ABL03413 ABL03413 standard; cDNA; 950 BP.

AC ABL03413;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4721.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX P-PSDB; ABB59310.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 genes from Drosophila and for elucidating cell signaling and cell-cell
 interactions.

PS Claim 1, SEQ ID NO 4721; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 capable of detecting 1000 or more genes from Drosophila. The invention is
 useful in developmental biology and in elucidating cell signalling and
 cell-cell interactions in higher eukaryotes for the development of
 insecticides, therapeutics and pharmaceutical drugs. The invention
 discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 sequences (AB01840-AB16175) and the encoded proteins (AB57737-
 AB372072). The sequence data for this patent did not form part of the
 printed specification, but was obtained in electronic format directly
 from WIPO at ftp.wipo.int/pub/published_pct_sequences

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match Best Local Similarity 82.0%; Score 16.4; DB 4; Length 950;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGCAACAGTACATGG 20

Db 634 TCGCAACAGTACATGG 651

RESULT 9

ABL16514/c

ID ABL16514 standard; DNA; 2901 BP.

XX ABL16514;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1015.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 1015; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB10511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2901 BP; 817 A; 687 C; 668 G; 729 T; 0 U; 0 Other;
XX
Query Match 82.0%; Score 16.4; DB 4; Length 2901;
Best Local Similarity 94.4%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 3 TCGCAACAGTACATGG 20
DB 2408 TCGCCACAGTACATGG 2391
XX
RESULT 10
AB103412/C
ID ABL03412 standard; cDNA; 3165 BP.
XX
XX ABL03412;
XX
DT 26-MAR-2002 (first entry)
XX
DE *Drosophila* melanogaster expressed polynucleotide SEQ ID NO 4718.
XX
XX *Drosophila*; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS *Drosophila* melanogaster.
XX
FN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
PD 23-MAR-2001; 2001WO-US009231.
XX
PE 23-MAR-2001; 2000US-0191637P.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW,
XX
DR WPI; 2001-6556860/75.
XX
DR P-PSDB; ABB59309.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 4718; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-

CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3165 BP; 866 A; 763 C; 700 G; 836 T; 0 U; 0 Other;
XX
Query Match 82.0%; Score 16.4; DB 4; Length 3165;
Best Local Similarity 94.4%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 3 TCGCAACAGTACATGG 20
DB 1317 TCGCCACAGTACATGG 1300
XX
RESULT 11
AAK54574
ID AAK54574 standard; cDNA; 182 BP.
XX
AC AAK54574;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen coding sequence #299.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
OS Homo sapiens.
XX
FN WO200164886-A2.
XX
XX 07-SEP-2001.
XX
PD 01-MAR-2001; 2001WO-US007272.
XX
PF 01-MAR-2001; 2000US-0186126P.
XX
PR 17-MAR-2000; 2000US-0190479P.
XX
PR 27-APR-2000; 2000US-0200545P.
XX
PR 28-APR-2000; 2000US-0200303P.
XX
PR 28-APR-2000; 2000US-0200779P.
XX
PR 01-MAY-2000; 2000US-0200999P.
XX
PR 04-MAY-2000; 2000US-0202084P.
XX
PR 22-MAY-2000; 2000US-0206201P.
XX
PR 14-JUL-2000; 2000US-0218958P.
XX
PR 03-AUG-2000; 2000US-0222903P.
XX
PR 04-AUG-2000; 2000US-0223416P.
XX
PR 07-AUG-2000; 2000US-0223378P.
XX
PA (CORI) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 411; 1252pp; English.
XX
PS The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma

XX SQ Sequence 182 BP; 49 A; 38 C; 48 G; 45 T; 0 U; 2 Other;
Query Match 79.0%; Score 15.8; DB 4; Length 182;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
100 AGTCGCAACAGTACATG 118
Do
RESULT 12
AAK54728
ID AAK54728 standard; cDNA; 182 BP.
XX AC AAK54728;
XX DT 13-NOV-2001 (first entry)
XX DE Human haematological malignancy-related antigen coding sequence #453.
XX DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX OS Homo sapiens.
XX PN WC200164886-A2.
XX PD 07-SEP-2001.
XX PF 01-MAR-2001; 2001WC-US007272.
XX PR 01-MAR-2000; 2000US-0186126P.
XX PR 17-MAR-2000; 2000US-0190479P.
XX PR 27-APR-2000; 2000US-0200545P.
XX PR 28-APR-2000; 2000US-0200303P.
XX PR 28-APR-2000; 2000US-0200779P.
XX PR 01-MAY-2000; 2000US-0200999P.
XX PR 04-MAY-2000; 2000US-0202084P.
XX PR 22-MAY-2000; 2000US-0206201P.
XX PR 14-JUL-2000; 2000US-0218950P.
XX PR 03-AUG-2000; 2000US-0222903P.
XX PR 04-AUG-2000; 2000US-0223416P.
XX PR 07-AUG-2000; 2000US-0223378P.
XX PA (CORI-) CORIXA CORP.
XX PI Gaiger A, Algate PA, Mannion J;
XX DR WPI; 2001-514842/56.
XX PT Compositions and methods for the detection of hematological malignancies,
XX PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX PS Claim 31; Page 451; 1252pp; English.
XX CC The present invention relates to compositions and methods for the
XX CC detection, diagnosis and therapy of haematological malignancies. The
XX CC present sequence is the coding sequence of a human haematological
XX CC malignancy related antigen. The methods of the present invention comprise
XX CC detecting the presence of haematological malignancy related antigen(s) in
XX CC a sample obtained from the patient (an increased level of the
XX CC polypeptide, compared to an unaffected individual, is indicative of an
XX CC increased risk). Haematological malignancies which can be treated using
XX CC the present invention are chronic lymphocytic leukaemia, lymphoma,
XX CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX CC cell non-Hodgkin's lymphoma
XX SQ Sequence 182 BP; 49 A; 38 C; 48 G; 45 T; 0 U; 2 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 182;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
100 AGTCGCAACAGTACATG 118
Db
RESULT 13
ADF79775/c
ID ADF79775 standard; DNA; 431 BP.
XX AC ADF79775;
XX DT 26-FEB-2004 (first entry)
XX DE Leukaemia-related DNA sequence #331.
XX DE Cytostatic; Gene therapy; leukaemia; ss.
XX OS Unidentified.
XX PN WC2003039443-A2.
XX PD 15-MAY-2003.
XX PF 04-NOV-2002; 2002WC-EP012303.
XX PR 05-NOV-2001; 2001EP-00126244.
XX PR 30-APR-2002; 2002EP-00009758.
XX PA (DEK-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
XX PA (HAE/) HAEFLACH T.
XX PA (SCH/) SCHROCH C.
XX PA (KERN/) KERN W.
XX PI Haefelach T, Schoch C, Kern W, Kohlmann A, Schmittger S, Dugas M;
XX PI Elis R, Broers B, Mergenthaler S;
XX DR WPI; 2003-505037/47.
XX PT Determining the subtype of leukemia cells and whether a patient sample
XX PT contains leukemia cells or other cells, useful for treating leukemia,
XX PT comprises determining the expression profile of a group of markers in a
XX PT patient sample.
XX PS Disclosure; SEQ ID NO 331; 2938pp; English.
XX CC The present invention relates to a method (M1) for determining the
XX CC subtype of leukemia cells and whether a patient sample contains
XX CC leukemia cells. The method comprises determining the expression profile
XX CC of a group of markers in a patient sample. The method is useful for
XX CC determining the presence of leukemia cells, its types or subtypes, and
XX CC for the preparation of a medicament for treating leukemia.
XX SQ Sequence 431 BP; 135 A; 104 C; 89 G; 103 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 10; Length 431;
Best Local Similarity 89.5%; Pred. No. 4e+02; 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
281 AGTCGCAACAGTACATG 263
Db
RESULT 14
ADF81342/c
ID ADF81342 standard; DNA; 440 BP.
XX AC ADF81342;

DT 11-FEB-2003 (first entry)
XX
DE S. pneumoniae type 4 strain coding region #881.
XX
KW Gene; ds: bacterial meningitis; pneumonia; sepsis; otitis media;
ear infection; antiinflammatory; antibacterial; immunostimulant;
XX auditory; respiratory; gene therapy; vaccine.
XX
OS Streptococcus pneumoniae; type 4 strain.
XX
PN W0200277021-A2.
XX
PD 03-OCT-2002.
XX
PF 27-MAR-2002; 2002W0-IB002163.
XX
PR 27-MAR-2001; 2001GB-00007658.
XX
PA (CHIR-) CHIRON SPA.
PI (GENO-) INST GENOMIC RES.
PI Masignani V, Tettelin H, Fraser C;
XX WPI: 2003-040579/03.
DR P-PSDB; ABU01306.
XX
PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
PS Claim 6; SEQ ID NO 1761; 56pp; English.
XX
XX The invention relates to a protein comprising or having at least 50%
XX identity to any of the 2469 amino acid sequences, identified in the
XX specification (available on a computer readable format), or its fragment,
XX expressed from 2469 of 2489 identified DNA coding regions from the
XX Streptococcus pneumoniae type 4 strain genomic sequence appearing as
XX AB556454. Also included are an antibody which binds one of the proteins,
XX treating a patient by administering the protein, DNA or antibody (in a
XX composition), a kit comprising first and second primers, which are the
XX nucleic acid cited above or fragments between nucleotides 8-100 of a
XX sequence not defined in the specification, for amplifying a target
XX sequence contained within a Streptococcus nucleic acid sequence, where
XX the first primer is substantially complementary to the target sequence
XX and the second primer is substantially complementary to the complement of
XX the target sequence, and where the parts of the primers having
XX substantial complementarity define the termini of the target sequence to
XX be amplified, assay comprising contacting a test compound with the
XX protein, and determining whether the test compound binds to the protein
XX and a Streptococcus pneumoniae bacterium, where one or more genes
XX encoding the proteins has been rendered inactive. The proteins, nucleic
XX acid molecules, antibody and compositions are useful as medicaments for
XX treating or preventing a disease or infection due to streptococcus
XX bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
XX media or ear infection. They are also useful in developing vaccines,
XX diagnostics and antibiotics. The methods are useful for identifying
XX immunodominant proteins. The present sequence is one of the 2469
XX identified coding regions from the genomic sequence. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences. (updated on 27-OCT-2003 to
XX standardise OS field)
XX
SQ Sequence 903 BP; 190 A; 205 C; 170 G; 338 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 10; Length 903;
XX
XX Best Local Similarity 89.5%; Pred. No. 4.3e+02;
XX
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 17
ADFO0300
ID ADFO0300 standard; DNA; 1281 BP.
XX
XX ADFO0300;
AC
XX 12-FEB-2004 (first entry)
DT
XX
DE Bacterial polynucleotide #585.
XX
XX Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant; gene; ds.
XX
OS Proteus mirabilis.
XX
XX US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
XX WPI: 2003-895291/82.
DR P-PSDB; ADF04472.
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
XX reagents for diagnosis of bacterial disease, as components of
XX antibacterial vaccines, as targets for antibacterial drugs, or as
XX biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 585; 870pp; English.
XX
XX The invention relates to new Proteus mirabilis polypeptides and
XX polynucleotides. The invention also relates to antibodies against the
XX polypeptides, methods for producing the polypeptides, a method of
XX generating vaccines for immunising an individual against P. mirabilis, a
XX method for evaluating a compound for the ability to bind a P. mirabilis
XX polypeptide and a method for screening test compounds for anti-bacterial
XX activity. The polypeptides and polynucleotides are useful as molecular
XX targets for diagnosing, preventing and treating pathological conditions
XX resulting from bacterial infection, as reagents for diagnosis of
XX bacterial diseases, as components of antibacterial vaccines, as targets
XX for antibacterial drugs or as bio-control agents for plants. This
XX sequence represents a Proteus mirabilis polynucleotide of the invention.
XX
SQ Sequence 1281 BP; 391 A; 235 C; 285 G; 370 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 10; Length 1281;
XX
XX Best Local Similarity 89.5%; Pred. No. 4.5e+02;
XX
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 18
ACA21176
ID ACA21176 standard; DNA; 1674 BP.
XX
XX ACA21176;
AC
XX 19-JUN-2003 (first entry)
DT
XX
DE Prokaryotic essential gene #2833.
XX

Antisense; ds; prokaryotic essential gene; cell proliferation;
 drug design; gene.
 Acinetobacter baumannii.
 WO200271183-A2.
 03-OCT-2002.
 21-MAR-2002; 2002WO-US009107.
 21-MAR-2001; 2001US-00815242.
 06-SEP-2001; 2001US-00948993.
 25-OCT-2001; 2001US-0342923P.
 08-FEB-2002; 2002US-00072851.
 06-MAR-2002; 2002US-0362699P.
 (ELITRA PHARM INC.
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 WPI; 2003-029926/02.
 P-PSDB; ABU17306.
 New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 Claim 14; SEQ ID NO 9046; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is one of the target
 prokaryotic essential genes. Note: The sequence data for this patent did
 not form part of the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Sequence 1674 BP; 476 A; 340 C; 410 G; 448 T; 0 U; 0 Other;
 Query Match 79.0%; Score 15.8; DB 8; Length 1674;
 Best Local Similarity 89.5%; Pred. No. 4.7e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 GTGGCAACAGTACATGG 20
 |||||
 473 GTGGCAACATTCATATGG 491

RESULT 19
 ADA29603
 ID ADA29603 standard; DNA; 1704 BP.
 XX
 AC ADA29603;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE DNA encoding Acinetobacter baumannii protein #890.
 XX
 KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 XX vaccine; plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00326352.
 XX
 PR 09-JUN-1998; 98US-0086701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 DR P-PSDB; ADA33729.
 XX
 PS Example; SEQ ID NO 890; 328pp; English.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. baumannii
 CC protein.
 CC
 XX
 SQ Sequence 1704 BP; 488 A; 342 C; 418 G; 456 T; 0 U; 0 Other;
 Query Match 79.0%; Score 15.8; DB 9; Length 1704;
 Best Local Similarity 89.5%; Pred. No. 4.7e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 GTGGCAACAGTACATGG 20
 |||||
 500 GTGGCAACATTCATATGG 518

RESULT 20
 ABR84220/C
 ID ABR84220 standard; cDNA; 2417 BP.
 XX
 AC ABR84220;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #791.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KM adult respiratory distress syndrome; inflammatory bowel disease;
 KM Crohn's disease; ulcerative colitis; periodontal disease;
 KM granulocyte activation; chronic inflammation; allergy.
 XX Homo sapiens.
 OS
 XX
 PN W0200228999-A2.
 PD
 XX 11-APR-2002.
 PF 03-OCT-2001; 2001WO-US030821.
 XX
 PR 03-OCT-2000; 2000US-0237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 PS
 XX Claim 1; SEQ ID NO 791; 114pp; English.
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GCA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and MS is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2417 BP; 788 A; 503 C; 521 G; 605 T; 0 U; 0 Other;
 Query Match 79.0%; Score 15.8; DB 6; Length 2417;
 Best Local Similarity 89.5%; Pred. No. 4.9e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 21
 ADO19070/C
 ID ADO19070 standard; cDNA; 2417 BP.
 XX
 AC ADO19070;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human PRO polynucleotide #4.
 XX
 KM Human, PRO, gene; ss; immune related disorder;
 KM systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KM juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KM vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KM autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KM renal disease; demyelinating disease; central nervous system;
 KM peripheral nervous system; demyelinating polyneuropathy;
 KM Guillain-Barre syndrome;
 KM chronic inflammatory demyelinating polyneuropathy.
 OS
 XX Homo sapiens.
 XX
 PN W02000403361-A2.
 PD 27-MAY-2004.
 PF 06-NOV-2003; 2003WO-US035268.
 XX
 PR 08-NOV-2002; 2002US-0425235P.
 XX
 PA (GENE-) GENENTECH INC.
 PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WJ, Wu TD;
 DR WPI; 2004-420067/39.
 XX
 DR P-PsDB; ADO19071.
 XX
 PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthropathy.
 PS
 XX Claim 1; SEQ ID NO 7; 1731pp; English.
 XX
 CC The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating polyneuropathy,
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.
 XX
 SQ Sequence 2417 BP; 788 A; 503 C; 521 G; 605 T; 0 U; 0 Other;
 Query Match 79.0%; Score 15.8; DB 12; Length 2417;
 Best Local Similarity 89.5%; Pred. No. 4.9e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATG 19
 DB 2101 AGCTGCAACAGTACATG 2083

RESULT 22
 ADO19457/C
 ID ADO19457 standard; cDNA; 2417 BP.

PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter UC, Adams M, Li FMD, Myers EW,
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB60264.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 7583; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB170511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2772 BP; 761 A; 684 C; 689 G; 638 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 4; Length 2772;
Best Local Similarity 89.5%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AGTCGCAACAGTACAAATG 19
Db 493 AATGCCAAACAGTACAAATG 511
XX
RESULT 25
ADB58306/c
ID ADB58306 standard; DNA; 4183 BP.
XX
AC ADB58306;
XX
DT 04-DEC-2003 (first entry)
XX
DE Toxicity-related gene, SEQ ID 3332.
XX
KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.
XX
XX Unidentified.
XX
XX WO2003064624-A2.
XX
XX 07-AUG-2003.
XX
XX 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-0060087.
XX
XX 15-MAR-2002; 2002US-0364045P.
XX
XX 15-MAR-2002; 2002US-0364055P.
XX
XX 30-DEC-2002; 2002US-0436643P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-689530/65.
XX
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to

PT the compound.
XX
XX Claim 1; SEQ ID NO 3332; 1156pp; English.
XX
XX
CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4183 BP; 1112 A; 1050 C; 926 G; 1095 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 10; Length 4183;
Best Local Similarity 89.5%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 2 GTTCGCAACAGTACAAATG 20
Db 3404 GTTCGCAACAGTACAAAGG 3386
XX
RESULT 26
ADB69202
ID ADB69202 standard; DNA; 5238 BP.
XX
AC ADB69202;
XX
DT 04-DEC-2003 (first entry)
XX
DE C. neoformans genomic DNA sequence SEQ ID NO:329.
XX
XX ds; gene; fungicide; gene therapy; infection.
XX
XX Cryptococcus neoformans.
XX
XX WO2003052076-A2.
XX
XX 26-JUN-2003.
XX
XX 17-DEC-2002; 2002WO-US040225.
XX
XX 17-DEC-2001; 2001US-0341261P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Zamudio C, Eroshkin AM;
XX
XX WPI; 2003-533017/50.
XX
XX P-PSDB; ADB70285.
XX
XX New nucleic acid, useful for preparing a composition for treating an
XX infection caused by Cryptococcus neoformans.
XX
XX Claim 3; SEQ ID NO 329; 136pp; English.
XX
XX The invention relates to a novel purified or isolated Cryptococcus
XX neoformans nucleic acid molecule comprising a sequence encoding a
XX polypeptide comprising a sequence not given in the specification. A
XX polynucleotide of the invention has fungicide activity, and may have a
XX use in gene therapy. The nucleic acid is useful for preparing a
XX composition for treating an infection caused by Cryptococcus neoformans.

CC The present sequence represents a C. neoformans sequence of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 5238 BP; 1147 A; 1479 C; 1211 G; 1400 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 10; Length 5238;

Best Local Similarity 89.5%; Pred. No. 5.3e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATG 19
|||||

DB 51 AGTCGCAACAGCGCATG 69

RESULT 27
AAVS2303/c
ID AAVS2303 standard; DNA; 8876 BP.

XX AAVS2303;

XX 23-OCT-1998 (first entry)

XX Streptococcus pneumoniae genome fragment SEQ ID NO:170.

XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;

XX computer readable medium; vaccine; pharmaceutical composition; de.

XX Streptococcus pneumoniae.

XX WO9818931-A2.

XX 07-MAY-1998.

XX 30-OCT-1997; 97WO-US019588.

XX 31-OCT-1996; 96US-0029960P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Dillon PJ, Rosen CA, Barash SC, Fannon M;

XX Dougherty BA;

XX WI; 1998-272225/24.

XX Claim 1; Page 1080-1085; 1409P; English.

XX The present invention describes a computer readable medium which has the
XX nucleotide sequences SEQ ID NO:1 to 391 (AAVS2134 to AAVS2524) recorded
XX on it, or a representative fragment or a sequence at least 95% identical
XX to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX to 391 (AAVS2134 to AAVS2524) are genomic fragments from Streptococcus
XX pneumoniae. The present invention also describes an isolated nucleic acid
XX molecule encoding a homologue of any of the fragments of the S. pneumoniae
XX genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX by a process comprising: (a) screening a genomic DNA library using as a
XX probe a target sequence defined by any of the sequences in SEQ ID NO:1 to
XX 391, identifying members of the library which contain sequences that
XX hybridize to the target sequence and isolating the nucleic acid molecules
XX from the members; or (b) isolating mRNA, DNA or cDNA produced from an
XX organism, amplifying nucleic acid molecules whose nucleotide sequence is
XX homologous to amplification primers derived from the fragment of the S.
XX pneumoniae genome to prime the amplification and isolating the amplified
XX sequences. The computer readable medium can be used in a computer-based
XX system for identifying fragments of the S. pneumoniae genome of
XX commercial importance, or expression modulating fragments of the S.
XX pneumoniae genome. Products from the present invention can be used in
XX diagnosis kits and assays, and pharmaceutical compositions and vaccines

CC for S. pneumoniae

XX SQ Sequence 8876 BP; 2663 A; 1840 C; 1505 G; 2867 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 2; Length 8876;

Best Local Similarity 89.5%; Pred. No. 5.7e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCGCAACAGTACATG 20
|||||

DB 4731 GTCGCAACAGCAATG 4713

RESULT 28
ABS56454_07
Continuation (8 of 22) of ABS56454 from base 700001 (Streptococcus pneumoniae type 4 str:

WP Sequence split into 22 fragments LOCUS ABS56454 Accession ABS56454

WP Fragment Name Begin End

WP ABS56454_00 1 11000

WP ABS56454_01 100001 21000

WP ABS56454_02 200001 31000

WP ABS56454_03 300001 41000

WP ABS56454_04 400001 51000

WP ABS56454_05 500001 61000

WP ABS56454_06 600001 71000

WP ABS56454_07 700001 81000

WP ABS56454_08 800001 91000

WP ABS56454_09 900001 101000

WP ABS56454_10 1000001 111000

WP ABS56454_11 1100001 121000

WP ABS56454_12 1200001 131000

WP ABS56454_13 1300001 141000

WP ABS56454_14 1400001 151000

WP ABS56454_15 1500001 161000

WP ABS56454_16 1600001 171000

WP ABS56454_17 1700001 181000

WP ABS56454_18 1800001 191000

WP ABS56454_19 1900001 201000

WP ABS56454_20 2000001 211000

WP ABS56454_21 2100001 2162598

Query Match 79.0%; Score 15.8; DB 10; Length 110000;
Best Local Similarity 89.5%; Pred. No. 7.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCGCAACAGTACATG 20
|||||

DB 52380 GTCGCAACAGCAATG 52398

RESULT 29
ADQ20017
ID ADQ20017 standard; DNA; 260160 BP.

XX ADQ20017;

XX 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 2837.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;

XX ds.

XX Homo sapiens.

XX WO2004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnick A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 2837; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 260160 BP; 68455 A; 56280 C; 57380 G; 78045 T; 0 U; 0 Other;
XX
QY Query Match 79.0%; Score 15.8; DB 12; Length 260160;
Best Local Similarity 89.5%; Pred. NO. 8.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 AGTGGCAACAGTACATG 19
152805 AATGGCAACAGTACATG 152823
XX
RESULT 30
ADB82581/C
ID ADB82581 standard; cDNA; 236 BP.
XX
AC ADB82581;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 893).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
OS
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX

PS Claim 1; SEQ ID NO 893; 188pp; English.
XX
CC This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 236 BP; 66 A; 41 C; 54 G; 49 T; 0 U; 26 Other;
XX
QY Query Match 77.0%; Score 15.4; DB 9; Length 236;
Best Local Similarity 94.1%; Pred. NO. 6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AGTGGCAACAGTACCA 17
202 AGTGGCAACAGTACCA 186
XX
DB ABO89640/C
ID ABO89640 standard; cDNA; 306 BP.
XX
AC ABO89640;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 896.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW ss.
XX
OS Homo sapiens.
OS
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US047349.
XX
PR 07-DEC-2000; 2000US-0254648P.
PR 13-MAR-2001; 2001US-0275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases.
XX
PS Claim 1; SEQ ID NO 896; 186pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung

CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC rabbits, horse or human). The polynucleotides and polypeptides are also
CC useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ8745-ABQ90015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequence
XX
SQ Sequence 306 BP; 83 A; 55 C; 76 G; 61 T; 0 U; 31 Other;
Query Match 77.0%; Score 15.4; DB 6; Length 306;
Best Local Similarity 94.1%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGTGCACAACAGTACAA 17
DB 272 AGTGCACAACAGTACCA 256
RESULT 32
ABV77453/C
ID ABV77453 standard; cDNA; 306 BP.
XX
AC ABV77453;
XX
DT 31-JAN-2003 (first entry)
XX
DE P. monodon full-length putative reproductive inhibiting hormone cDNA.
XX
KW Reproductive inhibiting hormone; spawning; penaeid; prawn; contraceptive;
KW antifertility; gene therapy; SGPVI; RH; PmSGPVI; gene; ss.
XX
OS Penaeus monodon.
XX
FH Key Location/Qualifiers
FT 1..306
FT CDS /*tag= b
FT /product= "Reproductive inhibiting hormone"
FT /note= "No stop codon given"
FT sig_peptide 1..81
FT /*tag= c
FT mat_peptide 82..306
FT /*tag= d
XX
XX WO200283717-A1.
XX
XX 24-OCT-2002.
XX
XX 11-APR-2002; 2002MO-AUD00466.
XX
XX 11-APR-2001; 2001AU-00004368.
XX
XX 16-MAY-2001; 2001AU-00005049.
XX
XX (AUMA-) AUSTRALIAN INST MARINE SCI.
XX
XX
XX PI Davey ML, Hall MR, Swan JA, Wilson KJ;
XX
XX WPI; 2003-092997/08.
XX
XX P-PSDB; ABP57824, ABP57825.
XX
XX New peptides from Penaeus monodon, useful for inhibiting spawning in
XX penaeid prawns, or for screening inhibitors or antagonists of the
XX reproductive inhibiting hormone, which are useful for inducing spawning
XX in Crustacea.
XX
XX
XX PS Claim 9; Page 6; 70pp; English.
XX
XX The invention relates to a novel isolated polypeptide from Penaeus
XX monodon, an active fragment of the polypeptide, or a peptide with
XX substantial sequence identity to the polypeptide, which serves to inhibit

CC spawning in penaeid prawns. The protein of the invention has
CC contraceptive and antifertility activity. The invention may have a use
CC in gene therapy. The protein is useful for inhibiting spawning in penaeid
CC prawns. This peptide is also useful for screening inhibitors of RH. The
CC antibody or RNAi is useful as an antagonist of a reproductive inhibiting
CC hormone, and is useful for inducing spawning in Crustacea. The present
CC sequence encodes the P. monodon putative reproductive inhibiting hormone
CC (RH), also referred to as PmSGPVI or SGPVI
XX
SQ Sequence 306 BP; 78 A; 60 C; 85 G; 83 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 8; Length 306;
Best Local Similarity 94.1%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GTCCGAACAGTACAT 18
DB 56 GTCCGAACAGTACAT 40
RESULT 33
ACH49767/C
ID ACH49767 standard; cDNA; 436 BP.
XX
AC ACH49767;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human leukocyte cDNA #1361.
XX
KW Human; ss; sequencing by hybridization; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STRC/) STRCHS-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX
XX PS Claim 1; SEQ ID NO 36979; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence

CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030073623

XX Sequence 436 BP; 140 A; 79 C; 86 G; 131 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 9; Length 436;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 4 CGCAACAGTACATAG 20
DB 23 CGCAACAGTACATAG 7

RESULT 34
ABZ73091
ID ABZ73091 standard; cDNA; 499 BP.

XX ABZ73091;

DT 10-APR-2003 (first entry)

DE Rice leaf EST, SEQ ID NO:29.

XX Rice; leaf; EST; expressed sequence tag; plant; biochip; DNA array;
KM cloning; detection; heterosis; hybrid vigour; transgene detection;
KM herbicide screening; pesticide screening; disease diagnosis;
KM medical research; agriculture; ss.

XX Oryza sativa.

OS CN1364936-A.

XX 21-AUG-2002.

PF 31-OCT-2001; 2001CN-00137672.

PR 31-OCT-2001; 2001CN-00137672.

PA (UYZH-) UNIV ZHEJIANG.

PI Li D, Dong H;

XX WPI; 2003-000550/01.

XX New rice leaf expression sequence labels and constituted biochip.

PS Claim 1; Page 20 (disclosure); 29pp; Chinese.

XX The invention relates to 50 novel ESTs (expressed sequence tags; ABZ73063
CC -ABZ73113) obtained from a rice leaf cDNA library. The invention also
CC relates to a biochip comprising these ESTs. The biochip of the invention
CC can be used in a variety of agricultural applications. It can be used in
CC the cloning of genes which confer useful characteristics in crop species,
CC to predict hybrid vigour (heterosis) at an early stage, in the detection
CC of transgenic agricultural produce, in screening for novel herbicides and
CC pesticides and in disease diagnosis. The biochip may also be used in
CC medical research. The present sequence represents a rice leaf EST of the
CC invention

XX Sequence 499 BP; 116 A; 128 C; 132 G; 123 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 8; Length 499;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 4 CGCAACAGTACATAG 20
DB 319 CGCAACAGTACATAG 335

RESULT 35
ABA60404
ID ABA60404 standard; DNA; 542 BP.

XX ABA60404;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #8709.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

XX WO200157277-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000669.

PF 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human fetal liver.

PS Claim 1; SEQ ID NO 8709; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human fetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WPI

CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 4 CGCAACAGTACATAG 20
DB 58 CGCAACAGTACATAG 74

RESULT 36
AA140288
ID AA140288 standard; DNA; 542 BP.

XX AA140288;

DT 17-OCT-2001 (first entry)

DE Probe #8974 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder; ss.

XX Homo sapiens.

```
XX XX WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 25; SEQ ID NO 8974; 654bp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
XX Query Match 77.0%; Score 15.4; DB 4; Length 542;
XX Best Local Similarity 94.1%; Pred. No. 6.6e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74
XX
RESULT 37
ABA28630
ID ABA28630 standard; DNA; 542 BP.
XX
XX ABA28630;
XX AC
XX DT 23-JAN-2002 (first entry)
XX DE Probe #7096 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
```

```
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 1; SEQ ID NO 7096; 530bp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
XX Query Match 77.0%; Score 15.4; DB 4; Length 542;
XX Best Local Similarity 94.1%; Pred. No. 6.6e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74
XX
RESULT 38
AAK34570
ID AAK34570 standard; DNA; 542 BP.
XX
XX AAK34570;
XX AC
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 9127.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 9127; 658bp + Sequence listing; English.
```

XX CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74
RESULT 39
AAK08682
ID AAK08682 standard; DNA; 542 BP.
XX
AC AAK08682;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 8673.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 8673; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74
RESULT 40
ABS34344
ID ABS34344 standard; DNA; 542 BP.
XX
XX ABS34344;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID No 9334.
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 9334; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 1309 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74

RESULT 41
 ABS09204 standard; DNA: 542 BP.
 ID ABS09204 standard; DNA: 542 BP.
 AC ABS09204;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human genome-derived single exon probe from lung SEQ ID NO 9195.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN MO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001MO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 1; SEQ ID NO 9195; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human

CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemorrhage, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this parent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at ftp.wpi.edu/pub/published_pcr_sequences
 XX
 SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
 XX
 QY
 DB 4 CGCAACAGTACATGG 20
 58 CGCAACAGTACATAG 74
 XX
 ID ABV77452 standard; CDNA: 818 BP.
 XX
 AC ABV77452;
 XX
 DT 31-JAN-2003 (first entry)
 XX
 DS P. monodon full-length putative reproductive inhibiting hormone cDNA.
 XX
 KW Reproductive inhibiting hormone; spawning; penaeid; prawn; contraceptive;
 KW antifertility; gene therapy; SGPV; RH; PMSGPV; gene; ss.
 XX
 OS Penaeus monodon.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..65
 FT CDS /*tag= a
 FT /*tag= b
 FT /*tag= c
 FT sig_peptide 66..146
 FT mat_peptide 147..371
 FT /*tag= d
 FT 3'UTR 375..818
 FT /*tag= e
 XX
 PN WO200283717-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 11-APR-2002; 2002MO-AU000466.
 XX
 PR 11-APR-2001; 2001AU-00004368.
 PR 16-MAY-2001; 2001AU-00005049.
 XX
 PA (AUMA-) AUSTRALIAN INST MARINE SCI.
 XX
 PI Davey ML, Hall MR, Swan JA, Wilson KJ;
 XX
 DR WPI; 2003-092997/08.
 DR P-PSDB; ABP57824, ABP57825.
 XX
 CC New peptides from Penaeus monodon, useful for inhibiting spawning in
 CC penaeid prawns, or for screening inhibitors or antagonists of the
 CC reproductive inhibiting hormone, which are useful for inducing spawning
 CC in Crustacea.

PS Claim 8; Page 5-6; 70pp; English.
XX
CC The invention relates to a novel isolated polypeptide from Penaeus
CC monodon, an active fragment of the polypeptide, or a peptide with
CC substantial sequence identity to the polypeptide, which serves to inhibit
CC spawning in penaeid prawns. The protein of the invention has
CC contraceptive and antifertility activity. The invention may have a use
CC in gene therapy. The protein is useful for inhibiting spawning in penaeid
CC prawns. This peptide is also useful for screening inhibitors of RH. The
CC antibody or RNAi is useful as an antagonist of a reproductive inhibiting
CC hormone, and is useful for inducing spawning in Crustacea. The present
CC sequence encodes the P. monodon putative reproductive inhibiting hormone
CC (RH), also referred to as PmsGPVI or SGPVI
XX
SQ Sequence 818 BP; 237 A; 163 C; 172 G; 246 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 8; Length 818;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GTGCGAAGAGTACAAAT 18
DB 121 GTGCGAAGAGTACAAAT 105
RESULT 43
AAST7840/C
ID AAST7840 standard; cDNA; 854 BP.
XX
AC AAST7840;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13644.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-FSDB: ABG13653.
XX
PT New isolated polynucleotide and encoded polypeptides; useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 13644; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 854 BP; 259 A; 165 C; 208 G; 222 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 5; Length 854;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACAAATG 19
DB 833 TCGCAACAGTAAATG 817
RESULT 44
ACA20264
ID ACA20264 standard; DNA; 1134 BP.
XX
AC ACA20264;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #1921.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Staphylococcus aureus.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR P-FSDB: ABUI6394.
XX
PT New antisense nucleic acids; useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 8134; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a gene in an operon required for
CC proliferation; (8) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (9)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (10) manufacturing an antibiotic; (11) profiling a
CC compound's activity; (12) a culture complicating strains in which the gene
CC product is overexpressed or underexpressed; (13) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (14) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the targets
CC of prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

Sequence: 1134 BP; 402 A; 174 C; 230 G; 328 T; 0 U; 0 Other;

Query Match	77.0%;	Score 15.4;	DB 8;	Length 1134;
Best Local Similarity	94.1%;	Pred. No. 7.2e+02;		
Matches 16; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```

QY      1 AGTCGCAACAGTACAA 17
          |||||
Db      389 AATCGCAACAGTACAA 405

```

RESULT 45
AAS62439
ID AAS62439 standard; cDNA; 1704 BP.

DT 14-FEB-2002 (first entry)

DE cdna sequence #226 encoding novel human secreted protein

KM Human secreted protein; hyperproliferative disorder; autoimmune disorder
KM immune deficiency disorder; blood disorder; inflammatory disorder;
KM infectious disorder; gene therapy; antimicrobial; hepatotropic;
KM immunosuppressive; antineoplastic; ss.

OS Homo sapiens.

PN WO200177291-A2.

PD 18-OCT-2001.

PF 29-MAR-2001; 2001WO-US010485.

PR 06-APR-2000; 2000US-0195604P.

PA (GEMY) GENETICS INST INC.

PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ,

PI Gulukota K, Graham JR;

DR WPI; 2002-010900/01.

PT New polynucleotides encoding secreted proteins useful for treating e.g. asthma, HIV and Crohn's disease.

PS Claim 1; Page 192-193; 391pp; English.

CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been

CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are useful
CC for treating diseases such as hyperproliferative disorders (e.g. cancer,
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis) and
CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC the invention are also useful in gene therapy. AA562214-AA562838
CC represent the cDNA sequences of the invention that encode for novel human
CC secreted proteins

SQ Sequence 1704 BP; 492 A; 397 C; 379 G; 435 T; 0 U; 1 Other;

Query Match	77.0%;	Score 15.4;	DB 6;	Length 1704;
Best Local Similarity	94.1%;	Pred. No. 7.5e+02;		
Matches 16;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

DY 4 CGCAACAGTACAATGG 20
|||
Db 519 CGCAACAGTACAATAG 535

RESULT 46
AAS51976
ID AAS51976 standard; DNA; 2826 BP.

DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #3993.

KM Antisense; ds; prokaryotic cellular proliferation gene; antibiotic
KM antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001

PF 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P.
DE 22 MAY 2000 0000UTS-0306848P.

PR 26-MAY-2000; 2000US-0207727P.

27-NOV-2000; 2000US-0253625P.

PR 16-FEB-2001; 2001US-0269308P.

PA (ELIT-) ELITRA PHARM INC.
VV

PI Haselbeck R, Ohlson KL, Zys
PI Yamamoto PT, Yui HH.

XX
DP
WBT: 3001-611495/70

DR P-PSDB; AA034117.
XX

PT New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids.

XX	claim 27: SEO ID NO
PS	

xx The invention relates to antisense inhibi
cc

CC their use in the discovery of novel antibiotics, the essential genes

Staphylococcus aureus, *Salmonella typhi*, *Klebsiella pneumoniae*,

CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2826 BP; 995 A; 422 C; 590 G; 819 T; 0 U; 0 Other;

XX AC Query Match 77.0%; Score 15.4; DB 4; Length 2826;
XX AC Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
XX AC Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAA 17
389 AATCGCAACAGTACAA 405

Db

RESULT 47
ACF73815
ID ACF73815 standard; DNA; 2844 BP.
XX ACF73815;
XX 20-NOV-2003 (first entry)
XX Staphylococcus aureus DNA #1495.
XX Staphylococcus aureus DNA #1495.
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target; gene; ds.
XX Staphylococcus aureus.
XX WO200294868-A2.
XX 28-NOV-2002.
XX 27-MAR-2002; 2002WO-IB002637.
XX 27-MAR-2001; 2001GB-00007661.
XX (CHIR-) CHIRON SPA.
XX Maignant V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX P-PSDB; ABM72255.
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX Claim 6; SEQ ID NO 2989; 49pp; English.
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX Sequence 2844 BP; 1002 A; 417 C; 591 G; 834 T; 0 U; 0 Other;

XX Query Match 77.0%; Score 15.4; DB 8; Length 2844;
XX AC Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
XX AC Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAA 17
389 AATCGCAACAGTACAA 405

Db

RESULT 48
AAS54776
ID AAS54776 standard; DNA; 2847 BP.
XX AAS54776;
XX 13-FEB-2002 (first entry)
XX Staphylococcus aureus DNA for cellular proliferation protein #1088.
XX Staphylococcus aureus DNA for cellular proliferation protein #1088.
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.
XX Staphylococcus aureus.
XX WO200170955-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0194078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207727P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0255262P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX P-PSDB; AAU36917.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX Claim 27; SEQ ID NO 8413; 51pp; English.
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX themselves and the discovery of novel antibiotics, the essential genes
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence encodes an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2847 BP; 1008 A; 424 C; 592 G; 823 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 2847;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCCAAACAGTACAA 17
Db 389 AATGCCAAACAGTACAA 405

RESULT 49

ACA20256
ID ACA20256 standard; DNA; 2847 BP.

ACA20256;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #1913.

XM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

OS Staphylococcus aureus.

PN MO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002W0-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (EIT-). ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

DR WPI: 2003-029926/02.

XX P-PSDB; ABU16386.

PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 8126; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a gene in an operon required for

CC proliferation; (8) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway; (9)

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-regulated gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2847 BP, 1005 A; 419 C; 592 G; 831 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 8; Length 2847;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCCAAACAGTACAA 17
Db 389 AATGCCAAACAGTACAA 405

RESULT 50
ABL29542/C
ID ABL29542 standard; DNA; 4971 BP.

ABL29542;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40099.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmacological; gene; ds.

XX Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001W0-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li FWD, Myers EW;

DR WPI: 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

XX Claim 1; SEQ ID NO 40099; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-

CC ABR12072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

CC Sequence 4971 BP, 1370 A; 1099 C; 1069 G; 1433 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 4971;
Best Local Similarity 94.1%; Pred. No. 8.5e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Mon Dec 6 12:24:36 2004

us-10-050-189a-7.rng

Page 31

Qy 4 CGCAACAGTACATGG 20
| | | | | | | | | |
Db 760 CGCAACAGGACATGG 744

Search completed: December 3, 2004, 02:24:51
Job time : 321.789 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 23:31:20 ; Search time 626.579 Seconds.
(without alignments)
1509.457 Million cell updates/sec

Title: US-10-050-189a-7
Perfect score: 20
Sequence: 1 agtcgaacacagtcacatg 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :

GenBdb1:*
1: gb_ba:*
2: gb_btg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	AX481360 Sequence
2	20	100.0	66479	6	AX676048 Sequence
3	20	100.0	78376	9	AL359692 Human DNA
4	17.4	87.0	110000	2	AC148981-2 Continuation (3 of
5	17.4	87.0	135276	2	AC073603 Mus muscu
6	17.4	87.0	151365	2	AC140732 Rattus no
7	17.4	87.0	202183	2	AC099737 Mus muscu
8	17.4	87.0	205681	2	AC119471 Rattus no
9	17.4	87.0	215196	2	AC073796 Mus muscu
10	17.4	87.0	229472	2	AC149052 Rattus no
11	17.4	87.0	229472	2	AC140735 Rattus no
12	17.4	87.0	243210	2	AC136831 Rattus no
13	17.4	87.0	243702	2	AC115331 Rattus no
14	17.4	87.0	243702	2	AC094556 Rattus no
15	16.8	84.0	10601	14	HHV6U1102 X92436 Human herpe
16	16.8	84.0	56884	2	AC104588 Homo sapi
17	16.8	84.0	71049	2	AC104586 Homo sapi
18	16.8	84.0	110000	2	AC112029 Continuation (3 of
19	16.8	84.0	110000	2	AC112029_3 Continuation (4 of

20	16.8	84.0	112404	8	OSJN00288	AL731639 Oryza sat
21	16.8	84.0	116585	10	AL139153	AL139153 Homo sapi
22	16.8	84.0	120311	10	AC006945	AC006945 Mus muscu
23	16.8	84.0	128294	8	AC135225	AC135225 Oryza sat
24	16.8	84.0	137580	8	AF377947	AF377947 Oryza sat
25	16.8	84.0	155848	9	AC104996	AC104996 Homo sapi
26	16.8	84.0	159321	14	HHV6ACNM	X83413 Human herpe
27	16.8	84.0	159693	2	OSJN00286	AL731641 Oryza sat
28	16.8	84.0	161573	14	AB021506	AB021506 Human her
29	16.8	84.0	162114	14	AF157706	AF157706 Human her
30	16.8	84.0	164107	9	AC104982	AC104982 Homo sapi
31	16.8	84.0	167273	10	AC083894	AC083894 Mus muscu
32	16.8	84.0	178546	2	AC129943	AC129943 Mus muscu
33	16.8	84.0	180212	2	AC140318	AC140318 Mus muscu
34	16.8	84.0	187725	10	AC147514	AC147514 Mus muscu
35	16.8	84.0	207091	10	AC079443	AC079443 Mus muscu
36	16.8	84.0	228916	2	AC132022	AC132022 Rattus no
37	16.8	84.0	22916	2	AC112824	AC112824 Rattus no
38	16.8	84.0	240243	2	AC120133	AC120133 Mus muscu
39	16.8	84.0	282567	2	AC099214	AC099214 Rattus no
40	16.8	84.0	321695	1	BX571870	BX571870 Photorhab
41	16.8	84.0	348505	1	TAE56651	TAE56651 Tricicum
42	16.4	82.0	925	8	CO574602	CO574602 Sequence
43	16.4	82.0	950	6	AY060813	AY060813 Drosophila
44	16.4	82.0	1143	3	CQ594254	CQ594254 Sequence
45	16.4	82.0	2901	6	CQ574601	CQ574601 Sequence
46	16.4	82.0	3165	6	AC017970	AC017970 Human DNA
47	16.4	82.0	89765	9	AL390057	AL390057 Human DNA
48	16.4	82.0	100578	9	BX649472	BX649472 Zebrafish
49	16.4	82.0	108623	5	AL365229	AL365229 Human DNA
50	16.4	82.0	121478	9	AL365222	AL365222 Human DNA
51	16.4	82.0	131111	10	AC140781	AC140781 Mus muscu
52	16.4	82.0	146687	9	AC110592	AC110592 Homo sapi
53	16.4	82.0	150053	8	AP003735	AP003735 Oryza sat
54	16.4	82.0	158144	2	CR550301	CR550301 Dario rer
55	16.4	82.0	174228	2	AC016105	AC016105 Homo sapi
56	16.4	82.0	188872	3	AC099009	AC099009 Drosophila
57	16.4	82.0	189125	2	CR392239	CR392239 Dario rer
58	16.4	82.0	191587	10	AC117610	AC117610 Mus muscu
59	16.4	82.0	192338	2	AC022221	AC022221 Homo sapi
60	16.4	82.0	193549	9	AC087286	AC087286 Homo sapi
61	16.4	82.0	194575	2	AC023140	AC023140 Homo sapi
62	16.4	82.0	196337	3	AC005894	AC005894 Drosophila
63	16.4	82.0	200626	8	AC109929	AC109929 Oryza sat
64	16.4	82.0	204616	2	AC132885	AC132885 Mus muscu
65	16.4	82.0	242172	3	AB003832	AB003832 Drosophila
66	16.4	82.0	244805	2	CR382321	CR382321 Dario rer
67	16.4	82.0	302963	2	CR394524	CR394524 Dario rer
68	16.4	82.0	321304	3	CEY105C5B	AL110479 Caenorhab
69	16.4	82.0	321304	2	AC087120	AC087120 Mus muscu
70	16.4	82.0	177688	10	AC127274	AC127274 Mus muscu
71	16.4	82.0	187484	2	AC119772	AC119772 Rattus no
72	16.4	82.0	210560	2	AC131578	AC131578 Mus muscu
73	16.4	82.0	245764	2	AC106238	AC106238 Rattus no
74	16.4	82.0	261998	2	AC107722	AC107722 Mus muscu
75	16.4	82.0	282009	2	AC131579	AC131579 Mus muscu
76	15.8	79.0	182	6	AX37477	AX374777 Sequence
77	15.8	79.0	182	6	AX37477	AX374777 Sequence
78	15.8	79.0	190	3	SUR5818	AX778174 Sequence
79	15.8	79.0	431	6	AX778174	AX778174 Sequence
80	15.8	79.0	440	6	AX779741	AX779741 Sequence
81	15.8	79.0	642	11	PM5D3B	AX686042 Penicilliu
82	15.8	79.0	903	6	AX568554	AX568554 Sequence
83	15.8	79.0	906	5	AR481547	AR481547 Sequence
84	15.8	79.0	955	5	BX322307	BX322307 Gallus ga
85	15.8	79.0	983	8	BX34761	BX34761 Gallus ga
86	15.8	79.0	1000	8	AK065851	AK065851 Oryza sat
87	15.8	79.0	1052	8	AK119532	AK119532 Oryza sat
88	15.8	79.0	1281	6	AR375579	AR375579 Sequence
89	15.8	79.0	1353	10	CR338752	CR338752 Gallus ga
90	15.8	79.0	1523	5	BC022458	BC022458 Mus muscu
91	15.8	79.0	1704	6	AR318340	AR318340 Sequence
92	15.8	79.0	1965	3	AF495359	AF495359 Drosophila

93	15.8	79.0	1966	3	AY128486	AY128486 Drosophila	166	15.8	79.0	166660	9	AC008073	AC008073 Homo sapi
94	15.8	79.0	2004	6	AX461258	AX461258 Sequence	167	15.8	79.0	168994	9	AC097480	AC097480 Homo sapi
95	15.8	79.0	2330	6	CG714512	CG714512 Sequence	168	15.8	79.0	169834	10	AC124430	AC124430 Mus muscu
96	15.8	79.0	2417	9	HSU10485	U10485 Human lymph	169	15.8	79.0	171089	9	AC006368	AC006368 Homo sapi
97	15.8	79.0	2716	6	AX781111	AX781111 Sequence	170	15.8	79.0	171577	10	AC106830	AC106830 Homo sapi
98	15.8	79.0	2772	6	CG576033	CG576033 Sequence	171	15.8	79.0	173519	9	AC145860	AC145860 Pan trogl
99	15.8	79.0	4183	10	RNGABAR1	X95579 R. norvegicu	172	15.8	79.0	173954	9	AC023108	AC023108 Homo sapi
100	15.8	79.0	8876	6	CQ789079	CQ789079 Sequence	173	15.8	79.0	173956	9	AC105383	AC105383 Homo sapi
101	15.8	79.0	8876	6	AR218938	AR218938 Sequence	174	15.8	79.0	174133	8	AF083031	AF083031 Guillard
102	15.8	79.0	8876	6	BD003850	BD003850 Polynucle	175	15.8	79.0	175552	2	AC119624	AC119624 Rattus no
103	15.8	79.0	10029	1	AE007396	AE007396 Streptococ	176	15.8	79.0	175552	2	AC091721	AC091721 Pan trogl
104	15.8	79.0	10615	1	AE008456	AE008456 Streptococ	177	15.8	79.0	175808	9	AC146040	AC146040 Homo sapi
105	15.8	79.0	10846	1	AE013594	AE013594 Methanosa	178	15.8	79.0	177716	5	AP001027	AP001027 Homo sapi
106	15.8	79.0	14399	8	SPAC1F12	269944 S. pombe chr	179	15.8	79.0	177074	5	BX539324	BX539324 Zebrafish
107	15.8	79.0	26333	9	CBRG17X04	AC084514 Caenorhab	180	15.8	79.0	178403	5	AL845283	AL845283 Mouse DNA
108	15.8	79.0	26713	9	CBRG17X04	AC136785 Homo sapi	181	15.8	79.0	178801	10	EX088688	EX088688 Homo sapi
109	15.8	79.0	37521	2	AC149354	AC149354 Plakospor	182	15.8	79.0	180565	2	AP002409	AP002409 Homo sapi
110	15.8	79.0	46077	2	AC137289	AC137289 Rattus no	183	15.8	79.0	182041	2	AC136153	AC136153 Rattus no
111	15.8	79.0	51424	5	BX276084	BX276084 Zebrafish	184	15.8	79.0	182228	10	AC036145	AC036145 Mus muscu
112	15.8	79.0	52204	2	AC015888	AC015888 Homo sapi	185	15.8	79.0	182712	9	AC016953	AC016953 Homo sapi
113	15.8	79.0	68191	2	AC136387	AC136387 Rattus no	186	15.8	79.0	183368	2	AC117878	AC117878 Rattus no
114	15.8	79.0	68548	9	HSU70984	U70984 Human Xp22	187	15.8	79.0	183602	2	AC024025	AC024025 Homo sapi
115	15.8	79.0	72652	8	NCEB912	AL451019 Neurospor	188	15.8	79.0	184781	2	CR388151	CR388151 Danio rer
116	15.8	79.0	74417	8	NCEB1D4	AL355928 Neurospor	189	15.8	79.0	184892	10	AC016017	AC016017 Mus muscu
117	15.8	79.0	82675	2	AC139375	AC139375 Mus muscu	190	15.8	79.0	186822	10	AL837505	AL837505 Mouse DNA
118	15.8	79.0	88366	2	AC104193	AC104193 Mus muscu	191	15.8	79.0	187972	2	AC146096	AC146096 Pan trogl
119	15.8	79.0	93225	2	AC021825	AC021825 Homo sapi	192	15.8	79.0	189458	9	AC091602	AC091602 Homo sapi
120	15.8	79.0	94038	8	AC004401	AC004401 Arabidops	193	15.8	79.0	190298	10	AC102676	AC102676 Mus muscu
121	15.8	79.0	107100	9	AC117377	AC117377 Homo sapi	194	15.8	79.0	195167	2	AC119829	AC119829 Mus muscu
122	15.8	79.0	110100	9	AC108405	AC108405 Mus muscu	195	15.8	79.0	195178	2	AC093471	AC093471 Mus muscu
123	15.8	79.0	110000	2	AC123189_4	Continuation (5 of	196	15.8	79.0	195187	5	EX957082	EX957082 Danio rer
124	15.8	79.0	112229	2	AC150129	AC150129 Gallus ga	197	15.8	79.0	195316	5	AL923082	AL923082 Zebrafish
125	15.8	79.0	112679	2	AC142219	AC142219 Rattus no	198	15.8	79.0	195667	10	AC023175	AC023175 Mus muscu
126	15.8	79.0	113006	8	AC006978	AC006978 Homo sapi	199	15.8	79.0	195901	10	AL844848	AL844848 Mouse DNA
127	15.8	79.0	113769	9	AP003799	AP003799 Oryza sat	200	15.8	79.0	199747	10	AL670024	AL670024 Mouse DNA
128	15.8	79.0	125937	9	AC092794	AC092794 Homo sapi	201	15.8	79.0	199908	2	AC111316	AC111316 Rattus no
129	15.8	79.0	125937	9	HS69M21	AL001735 Human DNA	202	15.8	79.0	200108	9	AP005271	AP005271 Homo sapi
130	15.8	79.0	126315	2	AP004116	AP004116 Oryza sat	203	15.8	79.0	202742	10	AL691493	AL691493 Mouse DNA
131	15.8	79.0	127673	2	H0201G08	AL442118 Oryza sat	204	15.8	79.0	203591	9	AC007156	AC007156 Homo sapi
132	15.8	79.0	130634	8	AP005476	AP005476 Oryza sat	205	15.8	79.0	206512	2	AC135435	AC135435 Rattus no
133	15.8	79.0	130952	8	AC146681	AC146681 Medicago	206	15.8	79.0	206877	2	AC102753	AC102753 Mus muscu
134	15.8	79.0	132017	8	AC108875	AC108875 Oryza sat	207	15.8	79.0	209021	10	AC098878	AC098878 Mus muscu
135	15.8	79.0	134426	8	AC092781	AC092781 Oryza sat	208	15.8	79.0	209561	2	CR383669	CR383669 Danio rer
136	15.8	79.0	136253	2	AC149939	AC149939 Strongylo	209	15.8	79.0	209643	10	AC079438	AC079438 Mus muscu
137	15.8	79.0	138600	2	AC150080	AC150080 Gallus ga	210	15.8	79.0	209705	2	AC007931	AC007931 Homo sapi
138	15.8	79.0	141266	2	BX571727	BX571727 Danio rer	211	15.8	79.0	212024	2	AC073390	AC073390 Mus muscu
139	15.8	79.0	144165	9	AC002137	AC002137 Homo sapi	212	15.8	79.0	212683	2	BX936317	BX936317 Danio rer
140	15.8	79.0	144767	8	CNS08C60	AL731884 Oryza sat	213	15.8	79.0	213033	5	AC131482	AC131482 Homo sapi
141	15.8	79.0	146243	2	AC150141	AC150141 Gallus ga	214	15.8	79.0	213989	2	AC145509	AC145509 Gasterost
142	15.8	79.0	146363	5	AL952423	AL952423 Homo sapi	215	15.8	79.0	216919	2	CR450833	CR450833 Danio rer
143	15.8	79.0	146415	5	AL929469	AL929469 Zebrafish	216	15.8	79.0	217249	10	AL663095	AL663095 Mouse DNA
144	15.8	79.0	146519	8	CNS08C9V	AL772417 Oryza sat	217	15.8	79.0	218724	2	AC025889	AC025889 Homo sapi
145	15.8	79.0	148042	2	AC115038	AC115038 Mus muscu	218	15.8	79.0	220758	10	AC117200	AC117200 Mus muscu
146	15.8	79.0	148543	9	AC087312	AC087312 Homo sapi	219	15.8	79.0	221384	2	AC141116	AC141116 Rattus no
147	15.8	79.0	148847	9	BS000227	BS000227 Pan trogl	220	15.8	79.0	222105	2	AC142172	AC142172 Rattus no
148	15.8	79.0	149472	2	AC140281	AC140281 Mus muscu	221	15.8	79.0	224597	2	AC114874	AC114874 Rattus no
149	15.8	79.0	149970	8	AC130607	AC130607 Oryza sat	222	15.8	79.0	224846	10	AC125327	AC125327 Mus muscu
150	15.8	79.0	149992	8	BX247886	BX247886 Zebrafish	223	15.8	79.0	228999	2	AC083885	AC083885 Homo sapi
151	15.8	79.0	150682	9	AC104842	AC104842 Homo sapi	224	15.8	79.0	229455	2	AC114352	AC114352 Rattus no
152	15.8	79.0	150872	9	AC103734	AC103734 Homo sapi	225	15.8	79.0	232282	2	AC094738	AC094738 Rattus no
153	15.8	79.0	151298	5	BX537274	BX537274 Zebrafish	226	15.8	79.0	232181	2	AC130063	AC130063 Rattus no
154	15.8	79.0	151436	2	AC150177	AC150177 Gallus ga	227	15.8	79.0	235426	2	AC128048	AC128048 Rattus no
155	15.8	79.0	152890	10	AC121962	AC121962 Mus muscu	228	15.8	79.0	235990	2	AC150044	AC150044 Gallus ga
156	15.8	79.0	153241	5	AC023510	AC023510 Zebrafish	229	15.8	79.0	238803	2	AC114836	AC114836 Rattus no
157	15.8	79.0	157614	5	BX677675	BX677675 Homo sapi	230	15.8	79.0	238830	2	AC102034	AC102034 Mus muscu
158	15.8	79.0	159240	10	AC021309	AC021309 Homo sapi	231	15.8	79.0	240337	2	AC109745	AC109745 Rattus no
159	15.8	79.0	160510	2	AC124368	AC124368 Mus muscu	232	15.8	79.0	240337	2	AC103275	AC103275 Rattus no
160	15.8	79.0	161892	2	AC136667	AC136667 Rattus no	233	15.8	79.0	240501	2	AC110364	AC110364 Rattus no
161	15.8	79.0	162974	2	AC137339	AC137339 Rattus no	234	15.8	79.0	243046	2	AC097995	AC097995 Rattus no
162	15.8	79.0	162974	2	AC150079	AC150079 Gallus ga	235	15.8	79.0	243872	2	AC095137	AC095137 Rattus no
163	15.8	79.0	163669	8	CR339041	CR339041 Danio rer	236	15.8	79.0	246996	2	CR3392340	CR3392340 Danio rer
164	15.8	79.0	164819	8	AC114983	AC114983 Oryza sat	237	15.8	79.0	248294	2	SPNEN1903	SPNEN1903 Streptococ
165	15.8	79.0	166590	2	AP001076	AP001076 Homo sapi	238	15.8	79.0	249280	2	AC111494	AC111494 Rattus no

C 239	15.8	79.0	253749	2	AC109560	Rattus no	C 312	15.4	77.0	132790	9	HS167A14	Z94721 Human DNA s
C 240	15.8	79.0	255836	2	AC130013	Rattus no	C 313	15.4	77.0	135726	2	CR394528	CR394528 Danto rer
C 241	15.8	79.0	255487	2	AC124756	Mus muscu	C 314	15.4	77.0	136240	3	AC117070	AC117070 Dictyoste
C 242	15.8	79.0	260160	9	AC130454	Homo sapi	C 315	15.4	77.0	139944	2	CR392012	CR392012 Danto rer
C 243	15.8	79.0	260193	2	AC127152	Rattus no	C 316	15.4	77.0	140881	5	BX539349	BX539349 Zebrafish
C 244	15.8	79.0	269096	2	AC112556	Rattus no	C 317	15.4	77.0	144022	5	BX085954	BX085954 Zebrafish
C 245	15.8	79.0	269096	5	BX004800	Rattus no	C 318	15.4	77.0	144131	10	AC110822	AC110822 Mus muscu
C 246	15.8	79.0	271519	2	AC020887	Mus muscu	C 319	15.4	77.0	145536	5	BX682233	BX682233 Zebrafish
C 247	15.8	79.0	272329	2	AL928724	Danio rer	C 320	15.4	77.0	146204	5	BX323080	BX323080 Zebrafish
C 248	15.8	79.0	274570	2	AC111994	Rattus no	C 321	15.4	77.0	146885	2	BX957322	BX957322 Danto rer
C 249	15.8	79.0	276724	2	AC110104	Rattus no	C 322	15.4	77.0	148076	2	CR385021	CR385021 Danto rer
C 250	15.8	79.0	278521	2	AC093715	Homo sapi	C 323	15.4	77.0	148491	2	CR626666	CR626666 Danto rer
C 251	15.8	79.0	278781	2	AC099235	Rattus no	C 324	15.4	77.0	148904	5	BX088527	BX088527 Zebrafish
C 252	15.8	79.0	290841	1	AE017322	Listeria	C 325	15.4	77.0	149122	2	BX511502	BX511502 Danto rer
C 253	15.8	79.0	296613	2	AC135877	Rattus no	C 326	15.4	77.0	149316	5	AL929019	AL929019 Zebrafish
C 254	15.8	79.0	341944	2	AC079158	Homo sapi	C 327	15.4	77.0	153929	2	AC011883	AC011883 Homo sapi
C 255	15.8	79.0	349980	6	AX571762	Sequence	C 328	15.4	77.0	155542	2	BX510929	BX510929 Danto rer
C 256	15.4	77.0	306	6	AX541649	Sequence	C 329	15.4	77.0	155820	2	BX546478	BX546478 Danto rer
C 257	15.4	77.0	506	11	BV136587	PA00042	C 330	15.4	77.0	156583	2	CR394535	CR394535 Danto rer
C 258	15.4	77.0	542	6	CQ100115	Sequence	C 331	15.4	77.0	156938	5	BX510305	BX510305 Zebrafish
C 259	15.4	77.0	542	6	CQ139105	Sequence	C 332	15.4	77.0	157344	5	BX545847	BX545847 Zebrafish
C 260	15.4	77.0	542	6	CQ175700	Sequence	C 333	15.4	77.0	157379	5	AL928946	AL928946 Zebrafish
C 261	15.4	77.0	542	6	CQ222495	Sequence	C 334	15.4	77.0	158387	2	BX470205	BX470205 Danto rer
C 262	15.4	77.0	542	6	CQ260448	Sequence	C 335	15.4	77.0	158673	5	BX248105	BX248105 Zebrafish
C 263	15.4	77.0	542	6	CQ298090	Sequence	C 336	15.4	77.0	159419	5	AC144823	AC144823 Danto rer
C 264	15.4	77.0	542	6	CQ334579	Sequence	C 337	15.4	77.0	159928	3	AC010565	AC010565 Drosophi
C 265	15.4	77.0	573	8	AB114650	Citrus un	C 338	15.4	77.0	161362	2	CR394555	CR394555 Danto rer
C 266	15.4	77.0	573	8	AB114658	Citrus si	C 339	15.4	77.0	161366	5	BX005382	BX005382 Zebrafish
C 267	15.4	77.0	573	8	AB114666	Citrus li	C 340	15.4	77.0	161518	2	BX272222	BX272222 Danto rer
C 268	15.4	77.0	715	3	AB054188	Panauus 1	C 341	15.4	77.0	161518	2	BX679660	BX679660 Danto rer
C 269	15.4	77.0	819	5	BC075123	Xenopus 1	C 342	15.4	77.0	161982	10	AC124356	AC124356 Mus muscu
C 270	15.4	77.0	965	3	AB054989	Panauus m	C 343	15.4	77.0	162019	5	BX005271	BX005271 Zebrafish
C 271	15.4	77.0	1713	8	AB072343	Citrus un	C 344	15.4	77.0	163121	5	AL928637	AL928637 Zebrafish
C 272	15.4	77.0	1907	8	CS1319762	Citrus si	C 345	15.4	77.0	163321	2	BX571969	BX571969 Danto rer
C 273	15.4	77.0	2245	8	AF372617	Citrus x	C 346	15.4	77.0	164093	2	CR376741	CR376741 Danto rer
C 274	15.4	77.0	2536	3	AY061826	Drosophi	C 347	15.4	77.0	164628	9	CN601DV7	AL135858 Human chr
C 275	15.4	77.0	2844	6	AX620026	Sequence	C 348	15.4	77.0	165022	9	AC012289	AC012289 Homo sapi
C 276	15.4	77.0	3184	5	BC060440	Xenopus 1	C 349	15.4	77.0	166396	2	AC023957	AC023957 Homo sapi
C 277	15.4	77.0	4971	6	CO613796	Sequence	C 350	15.4	77.0	166468	2	BX363658	BX363658 Danto rer
C 278	15.4	77.0	10883	8	SPU57841	Schistosom	C 351	15.4	77.0	167187	3	AC010108	AC010108 Drosophi
C 279	15.4	77.0	10999	1	AE002536	Neisseria	C 352	15.4	77.0	167743	2	BX322580	BX322580 Danto rer
C 280	15.4	77.0	13086	6	AF353898	Sequence	C 353	15.4	77.0	168834	2	CR457462	CR457462 Danto rer
C 281	15.4	77.0	15286	1	AF194079	Neisseria	C 354	15.4	77.0	169991	5	BX004781	BX004781 Zebrafish
C 282	15.4	77.0	15424	6	CO613766	Sequence	C 355	15.4	77.0	170665	5	BX072563	BX072563 Zebrafish
C 283	15.4	77.0	37077	5	BX927083	Sequence	C 356	15.4	77.0	171015	5	BX005395	BX005395 Zebrafish
C 284	15.4	77.0	38280	5	BX276119	Zebrafish	C 357	15.4	77.0	171830	5	AL772273	AL772273 Zebrafish
C 285	15.4	77.0	39210	3	UA1508	Cnemidomd	C 358	15.4	77.0	172367	2	BX572636	BX572636 Danto rer
C 286	15.4	77.0	53354	1	AC078935	Staphyloc	C 359	15.4	77.0	173127	9	AC021451	AC021451 Homo sapi
C 287	15.4	77.0	57214	2	AC019870	Drosophi	C 360	15.4	77.0	173219	2	BX697738	BX697738 Danto rer
C 288	15.4	77.0	59620	2	AC014068	Drosophi	C 361	15.4	77.0	174135	9	AL359637	AL359637 Human DNA
C 289	15.4	77.0	63378	2	AC113131	Homo sapi	C 362	15.4	77.0	174257	2	BX308755	BX308755 Danto rer
C 290	15.4	77.0	69037	2	AC081497	Drosophi	C 363	15.4	77.0	174286	2	CR405666	CR405666 Danto rer
C 291	15.4	77.0	77731	5	AC087254	Danio rer	C 364	15.4	77.0	174419	2	BX649525	BX649525 Danto rer
C 292	15.4	77.0	86950	2	AC149547	Medicago	C 365	15.4	77.0	176690	9	AC079148	AC079148 Homo sapi
C 293	15.4	77.0	93853	2	CR388160	Continuatio	C 366	15.4	77.0	176986	6	BX48511	BX48511 Zebrafish
C 294	15.4	77.0	95798	2	AL954721	Zebrafish	C 367	15.4	77.0	177080	5	BX323984	BX323984 Zebrafish
C 295	15.4	77.0	97774	5	AL591399	Zebrafish	C 368	15.4	77.0	177608	3	AC099553	AC099553 Drosophi
C 296	15.4	77.0	99681	5	AL713850	Zebrafish	C 369	15.4	77.0	178874	2	BX323590	BX323590 Danto rer
C 297	15.4	77.0	105510	5	BX005456	Zebrafish	C 370	15.4	77.0	179005	9	AC088507	AC088507 Homo sapi
C 298	15.4	77.0	105607	5	AC149582	Zebrafish	C 371	15.4	77.0	180137	2	AC023081	AC023081 Homo sapi
C 299	15.4	77.0	108530	2	BX247884	Zebrafish	C 372	15.4	77.0	180305	5	BX004768	BX004768 Zebrafish
C 300	15.4	77.0	110000	1	BX571856	Continuatio	C 373	15.4	77.0	180835	2	BX557237	BX557237 Danto rer
C 301	15.4	77.0	110000	1	BX571857	Continuatio	C 374	15.4	77.0	180861	2	CR391970	CR391970 Danto rer
C 302	15.4	77.0	110000	2	BX571857	Continuatio	C 375	15.4	77.0	181320	5	BX479922	BX479922 Zebrafish
C 303	15.4	77.0	110000	1	BX321891	Danio rer	C 376	15.4	77.0	182960	5	AC010017	AC010017 Drosophi
C 304	15.4	77.0	110000	2	BX321891	Danio rer	C 377	15.4	77.0	183026	5	BX323086	BX323086 Zebrafish
C 305	15.4	77.0	110000	2	BX308723	Danio rer	C 378	15.4	77.0	183715	2	CR626906	CR626906 Danto rer
C 306	15.4	77.0	110000	2	BX908723	Danio rer	C 379	15.4	77.0	184431	2	CR628331	CR628331 Danto rer
C 307	15.4	77.0	114149	5	BX470242	Zebrafish	C 380	15.4	77.0	185515	2	BX577362	BX577362 Danto rer
C 308	15.4	77.0	115448	5	BX255508	Zebrafish	C 381	15.4	77.0	186780	2	CR626943	CR626943 Danto rer
C 309	15.4	77.0	124324	5	BX294657	Zebrafish	C 382	15.4	77.0	187061	2	AC122564	AC122564 Mus muscu
C 310	15.4	77.0	128110	2	BX936634	Danio rer	C 383	15.4	77.0	188341	3	AC005719	AC005719 Drosophi
C 311	15.4	77.0	128110	2	BX936634	Danio rer	C 384	15.4	77.0	188529	2	BX927370	BX927370 Danto rer

C 385	15.4	77.0	188746	5	AL928594	AL928594 Zebrafish	C 458	15.4	77.0	349980	6	AX417043	AX417043 Sequence
C 386	15.4	77.0	190095	2	BX276112	BX276112 Danio rer	C 459	15.2	76.0	174	4	AY530078	AY530078 Solenodon
C 387	15.4	77.0	191663	2	BX649269	BX649269 Danio rer	C 460	15.2	76.0	245	11	BV080588	BV080588 sc1218_p6
C 388	15.4	77.0	193472	2	CR388077	CR388077 Danio rer	C 461	15.2	76.0	245	11	BV080591	BV080591 sc1218_p6
C 389	15.4	77.0	193505	2	CR628366	CR628366 Danio rer	C 462	15.2	76.0	246	11	BV080580	BV080580 sc1218_p5
C 390	15.4	77.0	195651	5	CR294114	CR294114 Zebrafish	C 463	15.2	76.0	246	11	BV080582	BV080582 sc1218_p6
C 391	15.4	77.0	198070	2	BX649326	BX649326 Danio rer	C 464	15.2	76.0	246	11	BV080586	BV080586 sc1218_p6
C 392	15.4	77.0	199540	10	AL805952	AL805952 Mouse DNA	C 465	15.2	76.0	246	11	BV080589	BV080589 sc1218_p6
C 393	15.4	77.0	200066	2	BX004997	BX004997 Danio rer	C 466	15.2	76.0	246	11	BV080590	BV080590 sc1218_p6
C 394	15.4	77.0	200937	2	BX510646	BX510646 Danio rer	C 467	15.2	76.0	246	11	BV080592	BV080592 sc1218_p6
C 395	15.4	77.0	201132	2	BX927146	BX927146 Danio rer	C 468	15.2	76.0	246	11	BV080593	BV080593 sc1218_p6
C 396	15.4	77.0	201188	5	AC145974	AC145974 Gallus ga	C 469	15.2	76.0	254	11	BV080584	BV080584 sc1218_p6
C 397	15.4	77.0	201816	5	BX248121	BX248121 Zebrafish	C 470	15.2	76.0	254	11	BV080585	BV080585 sc1218_p6
C 398	15.4	77.0	203371	2	AC139623	AC139623 Danio rer	C 471	15.2	76.0	254	11	BV080587	BV080587 sc1218_p6
C 399	15.4	77.0	203944	2	BX537255	BX537255 Danio rer	C 472	15.2	76.0	342	6	AX918716	AX918716 Sequence
C 400	15.4	77.0	205102	2	CR336225	CR336225 Danio rer	C 473	15.2	76.0	342	6	BD054249	BD054249 Sequence
C 401	15.4	77.0	205289	5	AL954337	AL954337 Zebrafish	C 474	15.2	76.0	344	11	BV108269	BV108269 P2A01791
C 402	15.4	77.0	205802	5	AL772132	AL772132 Zebrafish	C 475	15.2	76.0	347	11	BV108277	BV108277 P2A01791
C 403	15.4	77.0	207146	2	CR388142	CR388142 Danio rer	C 476	15.2	76.0	349	11	BV108265	BV108265 P2A01791
C 404	15.4	77.0	207701	2	BX545856	BX545856 Danio rer	C 477	15.2	76.0	357	4	AF271977	AF271977 Bos tauru
C 405	15.4	77.0	208075	2	BX573363	BX573363 Danio rer	C 478	15.2	76.0	368	11	BV108271	BV108271 P2A01791
C 406	15.4	77.0	208770	2	CR398135	CR398135 Danio rer	C 479	15.2	76.0	372	11	BV108274	BV108274 P2A01791
C 407	15.4	77.0	209782	10	AL603905	AL603905 Mouse DNA	C 480	15.2	76.0	378	11	BV108272	BV108272 P2A01791
C 408	15.4	77.0	210355	2	BX511081	BX511081 Danio rer	C 481	15.2	76.0	380	11	BV108276	BV108276 P2A01791
C 409	15.4	77.0	210630	5	AL928276	AL928276 Zebrafish	C 482	15.2	76.0	384	11	BV108270	BV108270 P2A01791
C 410	15.4	77.0	212049	5	BX005484	BX005484 Zebrafish	C 483	15.2	76.0	385	11	BV108264	BV108264 P2A01791
C 411	15.4	77.0	212160	5	BX323453	BX323453 Zebrafish	C 484	15.2	76.0	385	11	BV108266	BV108266 P2A01791
C 412	15.4	77.0	212553	2	AC099581	AC099581 Mus muscu	C 485	15.2	76.0	385	11	BV108267	BV108267 P2A01791
C 413	15.4	77.0	216389	2	BX640518	BX640518 Danio rer	C 486	15.2	76.0	385	11	BV108268	BV108268 P2A01791
C 414	15.4	77.0	216615	10	AC107850	AC107850 Mus muscu	C 487	15.2	76.0	393	11	BV108275	BV108275 P2A01791
C 415	15.4	77.0	217582	5	BX005022	BX005022 Zebrafish	C 488	15.2	76.0	396	5	AY582107	AY582107 Leptocott
C 416	15.4	77.0	222751	2	CR538727	CR538727 Danio rer	C 489	15.2	76.0	513	6	AX818942	AX818942 Sequence
C 417	15.4	77.0	223880	2	AC103105	AC103105 Rattus no	C 490	15.2	76.0	513	6	AX825972	AX825972 Sequence
C 418	15.4	77.0	224218	2	CR556716	CR556716 Danio rer	C 491	15.2	76.0	513	8	AY557727	AY557727 Saccharom
C 419	15.4	77.0	225257	2	BX088574	BX088574 Danio rer	C 492	15.2	76.0	584	11	BV038129	BV038129 S212P6970
C 420	15.4	77.0	227706	5	BX571720	BX571720 Zebrafish	C 493	15.2	76.0	756	6	AR319714	AR319714 Sequence
C 421	15.4	77.0	227892	5	BX004471	BX004471 Zebrafish	C 494	15.2	76.0	883	8	BT013102	BT013102 Lycopersi
C 422	15.4	77.0	228584	2	CR392040	CR392040 Danio rer	C 495	15.2	76.0	1010	3	AF044817	AF044817 Papilio p
C 423	15.4	77.0	228902	2	AC098116	AC098116 Rattus no	C 496	15.2	76.0	1010	3	AF044819	AF044819 Papilio m
C 424	15.4	77.0	231484	2	AC096379	AC096379 Rattus no	C 497	15.2	76.0	1010	3	AF044820	AF044820 Papilio t
C 425	15.4	77.0	233451	2	AC103060	AC103060 Rattus no	C 498	15.2	76.0	1010	3	AF044830	AF044830 Papilio h
C 426	15.4	77.0	233058	5	BX640463	BX640463 Zebrafish	C 499	15.2	76.0	1045	3	AY457621	AY457621 Papilio m
C 427	15.4	77.0	234274	2	AC148857	AC148857 Oryzomur	C 500	15.2	76.0	1121	10	BC048519	BC048519 Mus muscu
C 428	15.4	77.0	235351	2	AC130518	AC130518 Rattus no	C 501	15.2	76.0	1220	3	AY457606	AY457606 Papilio c
C 429	15.4	77.0	235897	5	AC095685	AC095685 Rattus no	C 502	15.2	76.0	1240	3	AF044828	AF044828 Papilio m
C 430	15.4	77.0	238499	5	BX005043	BX005043 Zebrafish	C 503	15.2	76.0	1276	5	CR523553	CR523553 Gallus ga
C 431	15.4	77.0	238646	2	AC099463	AC099463 Rattus no	C 504	15.2	76.0	1278	6	AX848366	AX848366 Sequence
C 432	15.4	77.0	240539	2	BX571803	BX571803 Danio rer	C 505	15.2	76.0	1278	6	AX141387	AX141387 Sequence
C 433	15.4	77.0	240834	2	BX571803	BX571803 Danio rer	C 506	15.2	76.0	1349	6	AX685942	AX685942 Sequence
C 434	15.4	77.0	240834	2	BX095193	BX095193 Danio rer	C 507	15.2	76.0	1349	6	AX685942	AX685942 Sequence
C 435	15.4	77.0	243235	2	CR318614	CR318614 Danio rer	C 508	15.2	76.0	1643	5	BX934056	BX934056 Strongylo
C 436	15.4	77.0	246998	2	CR450707	CR450707 Danio rer	C 509	15.2	76.0	1702	3	AF248864	AF248864 Drosophi
C 437	15.4	77.0	248738	5	BX284112	BX284112 Danio rer	C 510	15.2	76.0	1905	3	AT047577	AT047577 Drosophi
C 438	15.4	77.0	249278	5	BX511172	BX511172 Zebrafish	C 511	15.2	76.0	1924	3	AF215861	AF215861 Drosophi
C 439	15.4	77.0	257502	5	BX005342	BX005342 Zebrafish	C 512	15.2	76.0	1941	3	MSF287537	MSF287537 Mesometra
C 440	15.4	77.0	258640	2	CR382325	CR382325 Danio rer	C 513	15.2	76.0	1950	8	AY087725	AY087725 Xenopus
C 441	15.4	77.0	259889	2	AC125704	AC125704 Rattus no	C 514	15.2	76.0	2003	5	AF369901	AF369901 Xenopus
C 442	15.4	77.0	260050	1	AL596166	AL596166 Listeria	C 515	15.2	76.0	2077	10	BC048839	BC048839 Mus muscu
C 443	15.4	77.0	261272	2	CR388059	CR388059 Danio rer	C 516	15.2	76.0	2089	5	BC048038	BC048038 Danio rer
C 444	15.4	77.0	266261	2	AE003664	AE003664 Drosophi	C 517	15.2	76.0	2248	5	AT051787	AT051787 Glycydus
C 445	15.4	77.0	266607	2	BX571847	BX571847 Danio rer	C 518	15.2	76.0	2334	6	AX934444	AX934444 Sequence
C 446	15.4	77.0	267191	2	AC132043	AC132043 Rattus no	C 519	15.2	76.0	2385	6	CQ594468	CQ594468 Sequence
C 447	15.4	77.0	282115	4	AE003545	AE003545 Drosophi	C 520	15.2	76.0	2525	10	BC057991	BC057991 Mus muscu
C 448	15.4	77.0	284624	2	BX511034	BX511034 Zebrafish	C 521	15.2	76.0	2609	3	AY548906	AY548906 Antonomsp
C 449	15.4	77.0	280150	1	AP004824	AP004824 Staphyloc	C 522	15.2	76.0	2699	9	HSTPM378	HSTPM378 Mus muscu
C 450	15.4	77.0	286107	2	BX545912	BX545912 Danio rer	C 523	15.2	76.0	2734	10	BC067200	BC067200 Mus muscu
C 451	15.4	77.0	299050	1	AP003131	AP003131 Staphyloc	C 524	15.2	76.0	2869	1	AF290090	AF290090 Staphyloc
C 452	15.4	77.0	300409	1	AE016755	AE016755 Escherich	C 525	15.2	76.0	2955	10	BC050127	BC050127 Mus muscu
C 453	15.4	77.0	314906	4	AE003594	AE003594 Drosophi	C 526	15.2	76.0	3007	2	AC015213	AC015213 Drosophi
C 454	15.4	77.0	326301	1	NMA622491	NMA622491 Neisseria	C 527	15.2	76.0	3044	10	AY032766	AY032766 Mus muscu
C 455	15.4	77.0	348477	1	AP003360	AP003360 Staphyloc	C 528	15.2	76.0	3172	1	AF269836	AF269836 Staphyloc
C 456	15.4	77.0	349980	6	AX044034	AX044034 Sequence	C 529	15.2	76.0	3172	1	AF270274	AF270274 Staphyloc
C 457	15.4	77.0	349980	6	AX417039	AX417039 Sequence	C 530	15.2	76.0	3172	6	AR485790	AR485790 Sequence

531	15.2	76.0	3172	6	AR468228	AR468228 Sequence	604	15.2	76.0	95637	8	EX42594	EX42594 Neurospor
532	15.2	76.0	3172	6	AX145154	AX145154 Sequence	605	15.2	76.0	98613	2	AC139630	AC139630 Takifugu
533	15.2	76.0	3172	6	AX145352	AX145352 Sequence	606	15.2	76.0	103041	9	AC093384	AC093384 Homo sapi
534	15.2	76.0	3138	8	AF302666	AF302666 Arabidops	607	15.2	76.0	103245	2	AC108382	AC108382 Pan trogl
535	15.2	76.0	3586	8	SCU07938	SCU07938 Saccharomyc	608	15.2	76.0	107113	3	AC142559	AC142559 Takifugu
536	15.2	76.0	3827	8	SCYN243W	271519 S.cerevisia	609	15.2	76.0	108688	8	AC105733	AC105733 Oryza sat
537	15.2	76.0	4206	3	DVZFP	227444 D.virilis m	610	15.2	76.0	110000	1	CR543861.22	CR543861.22
538	15.2	76.0	4530	10	AK129249	AK129249 Mus muscu	611	15.2	76.0	110000	1	CR543861.24	CR543861.24
539	15.2	76.0	4641	8	YSCMOP26	LI2352 Saccharomyc	612	15.2	76.0	110000	2	AC091338-2	AC091338-2
540	15.2	76.0	5130	6	CQ594467	CQ594467 Sequence	613	15.2	76.0	110000	2	AC117128-1	AC117128-1
541	15.2	76.0	5133	6	CQ583835	CQ583835 Sequence	614	15.2	76.0	110000	2	AC129112-0	AC129112-0
542	15.2	76.0	5323	5	AB074891	AB074891 Oryza sat	615	15.2	76.0	110000	2	AC138524-1	AC138524-1
543	15.2	76.0	5751	3	AF248863	AF248863 Strongylo	616	15.2	76.0	110000	2	AL137126-0	AL137126-0
544	15.2	76.0	5919	8	SCYN242W	271518 S.cerevisia	617	15.2	76.0	110000	2	AL451003-0	AL451003-0
545	15.2	76.0	5966	14	SYNCP	DI2517 Strawberry	618	15.2	76.0	110000	2	AL583837-2	AL583837-2
546	15.2	76.0	6187	8	SLA310657	AL310657 Silene la	619	15.2	76.0	110000	2	AL713976-1	AL713976-1
547	15.2	76.0	8382	6	CQ575612	CQ575612 Sequence	620	15.2	76.0	110000	2	AL929691-10	AL929691-10
548	15.2	76.0	10029	1	AE013415	AE013415 Methanosa	621	15.2	76.0	110000	2	BR294366-1	BR294366-1
549	15.2	76.0	11011	1	AE007252	AE007252 Sinorhizo	622	15.2	76.0	110000	2	BR294366-2	BR294366-2
550	15.2	76.0	11124	1	AE009714	AE009714 Brucella	623	15.2	76.0	110000	8	CR382132-27	CR382132-27
551	15.2	76.0	11565	1	AE010682	AE010682 Methanosa	624	15.2	76.0	110000	8	CR382132-20	CR382132-20
552	15.2	76.0	12311	8	AF004559	AF004559 Venturia	625	15.2	76.0	110000	9	HS46345-1	HS46345-1
553	15.2	76.0	12461	1	AE015834	AE015834 Shewanell	626	15.2	76.0	110000	10	AE014174-0	AE014174-0
554	15.2	76.0	12489	1	AE014544	AE014544 Brucella	627	15.2	76.0	111275	2	AC139876	AC139876
555	15.2	76.0	12500	1	AE015261	AE015261 Treponema	628	15.2	76.0	111375	2	AC073878	AC073878 Homo sapi
556	15.2	76.0	12500	1	AE015262	AE015262 Shigella f	629	15.2	76.0	115174	9	AC106628	AC106628 Homo sapi
557	15.2	76.0	13123	1	AE003872	AE003872 Xylella f	630	15.2	76.0	117695	2	AC137831	AC137831 Medicago
558	15.2	76.0	13852	10	AB026490	AB026490 Mus muscu	631	15.2	76.0	119504	5	BL448306	BL448306 Arabidops
559	15.2	76.0	15280	4	CEH31820	AL021066 Caenorhab	632	15.2	76.0	120466	9	AL449104	AL449104 Human DNA
560	15.2	76.0	16829	4	RUMTEBNOX	X97336 Rhinoceros	633	15.2	76.0	124587	2	AC141185	AC141185 Rattus no
561	15.2	76.0	18155	6	AX429460	AX429460 Sequence	634	15.2	76.0	125150	2	AC005450	AC005450 Drosophi
562	15.2	76.0	22062	6	AC019898	AC019898 Drosophi	635	15.2	76.0	125803	8	ATP22X18	ATP22X18 Homo sapi
563	15.2	76.0	22253	8	SC9320X	Z70202 S.cerevisia	636	15.2	76.0	126039	9	AC012665	AC012665 Homo sapi
564	15.2	76.0	22947	1	AE008776	AE008776 Salmonella	637	15.2	76.0	127347	10	AC102818	AC102818 Mus muscu
565	15.2	76.0	24000	8	SC9320A	Z68329 S.cerevisia	638	15.2	76.0	127395	7	AY176327	AY176327 Staphyloc
566	15.2	76.0	26568	9	HSV755G1	Z68329 Human DNA s	639	15.2	76.0	128809	9	AC021648	AC021648 Homo sapi
567	15.2	76.0	28534	9	EX572102	EX572102 Human DNA	640	15.2	76.0	129800	9	AL357353	AL357353 Human DNA
568	15.2	76.0	31180	3	DMC155E2	AL024485 Drosophi	641	15.2	76.0	130325	9	AC099566	AC099566 Homo sapi
569	15.2	76.0	32158	3	CEM04G3	Z68014 Caenorhabd	642	15.2	76.0	131815	9	AL627442	AL627442 Human DNA
570	15.2	76.0	32363	9	AC114792	AC114792 Homo sapi	643	15.2	76.0	131891	9	AC008131	AC008131 Homo sapi
571	15.2	76.0	33171	9	AC146820	AC146820 Homo sapi	644	15.2	76.0	132675	5	BX324182	BX324182 Zebrafish
572	15.2	76.0	33274	3	CEFS2B5	Z75541 Caenorhabd	645	15.2	76.0	132891	2	BX088564	BX088564 Danto rer
573	15.2	76.0	33855	8	SCCX733K	Z69381 S.cerevisia	646	15.2	76.0	132969	9	AC002403	AC002403 Human Chr
574	15.2	76.0	39800	2	AL513223	AL513223 Drosophi	647	15.2	76.0	133915	9	AC124055	AC124055 Homo sapi
575	15.2	76.0	40280	8	CEBRI.7N15	AC084515 Caenorhab	648	15.2	76.0	134157	2	AC141344	AC141344 Rattus no
576	15.2	76.0	40438	8	SPU4.410	U41410 Schizosacch	649	15.2	76.0	134325	2	AC141344	AC141344 Pan trogl
577	15.2	76.0	40478	3	U29614	AC068077 Homo sapi	650	15.2	76.0	134708	2	AC147046	AC147046 Homo sapi
578	15.2	76.0	40478	3	U29614	U29614 Caenorhabd	651	15.2	76.0	134708	9	AC137697	AC137697 Homo sapi
579	15.2	76.0	40545	8	AL500524	AL500524 Human DNA	652	15.2	76.0	135245	8	AL136092	AL136092 Human DNA
580	15.2	76.0	40810	8	SPBC1773	AL033389 S.pombe c	653	15.2	76.0	136150	8	AP002485	AP002485 Oryza sat
581	15.2	76.0	46279	2	AC014815	AC014815 Drosophi	654	15.2	76.0	136150	8	AP002485	AP002485 Oryza sat
582	15.2	76.0	49122	2	AC108385	AC108385 Pan trogl	655	15.2	76.0	136515	9	AL499609	AL499609 Human DNA
583	15.2	76.0	50089	3	AC005641	AC005641 Drosophi	656	15.2	76.0	136531	9	HS1050C22	HS1050C22 Oryza sat
584	15.2	76.0	51811	2	AL100162	AL100162 Mus muscu	657	15.2	76.0	137362	8	AP002820	AP002820 Oryza sat
585	15.2	76.0	54625	2	AL159979	AL159979 Human DNA	658	15.2	76.0	137432	8	AP002538	AP002538 Oryza sat
586	15.2	76.0	58320	2	AC133850	AC133850 Rattus no	659	15.2	76.0	138115	2	AC142040	AC142040 Rattus no
587	15.2	76.0	59492	2	EX908805	EX908805 Human DNA	660	15.2	76.0	140028	9	AC104658	AC104658 Homo sapi
588	15.2	76.0	64227	2	AC100013	AC100013 Mus muscu	661	15.2	76.0	140356	9	AC007528	AC007528 Genomic s
589	15.2	76.0	64430	2	AC083772	AC083772 Homo sapi	662	15.2	76.0	140619	9	AC007285	AC007285 Homo sapi
590	15.2	76.0	67435	2	AC010479	AC010479 Homo sapi	663	15.2	76.0	141396	2	AC091709	AC091709 Rattus no
591	15.2	76.0	72354	9	AL357123	AL357123 Human DNA	664	15.2	76.0	141973	2	AC099530	AC099530 Felis cat
592	15.2	76.0	73198	9	HSBD12713	AL096793 Human DNA	665	15.2	76.0	144062	5	AL954190	AL954190 Zebrafish
593	15.2	76.0	76336	5	AC023249	AC023249 Homo sapi	666	15.2	76.0	145704	3	AE002656	AE002656 Drosophi
594	15.2	76.0	80236	5	AL355138	AL355138 Human DNA	667	15.2	76.0	145880	9	HS302D9	HS302D9 Human DNA s
595	15.2	76.0	80804	10	EX510908	EX510908 Mouse DNA	668	15.2	76.0	147851	2	AC141379	AC141379 Rattus no
596	15.2	76.0	81874	5	EX323864	EX323864 Zebrafish	669	15.2	76.0	147875	9	AC067582	AC067582 Homo sapi
597	15.2	76.0	82291	9	AC005376	AC005376 Homo sapi	670	15.2	76.0	148819	9	AC006996	AC006996 Homo sapi
598	15.2	76.0	86564	1	AE008919	AE008919 Unculture	671	15.2	76.0	149200	9	AL359645	AL359645 Human DNA
599	15.2	76.0	86966	1	AC108720	AC108720 Homo sapi	672	15.2	76.0	149427	2	AC121385	AC121385 Rattus no
600	15.2	76.0	87428	3	CEY64G10A	AL110498 Caenorhab	673	15.2	76.0	149795	10	AC125173	AC125173 Mus muscu
601	15.2	76.0	87476	2	AC138695	AC138695 Homo sapi	674	15.2	76.0	150894	8	D90904	D90904 Synchocyst
602	15.2	76.0	89017	2	AC108387	AC108387 Pan trogl	675	15.2	76.0	151405	8	AP004261	AP004261 Oryza sat
603	15.2	76.0	91552	6	AX429455	AX429455 Sequence	676	15.2	76.0	151640	10	AC122053	AC122053 Mus muscu

677	15.2	76.0	151748	9	AC147408	Pan trogl	750	15.2	76.0	173518	10	AL772358	AL772358 Mouse DNA
678	15.2	76.0	151992	9	AC009939	Homo sapi	751	15.2	76.0	173558	9	AC093500	AC093500 Drosophila
679	15.2	76.0	152653	2	AC139129	Mus muscu	752	15.2	76.0	174223	2	AC128776	AC128776 Rattus no
680	15.2	76.0	152914	8	CNS06CAR	ALB37528	753	15.2	76.0	175285	3	AC008259	AC008259 Drosophila
681	15.2	76.0	153053	8	AC119281	Oryza sat	754	15.2	76.0	175485	3	AC147539	AC147539 Pan trogl
682	15.2	76.0	153094	9	CNS05TCH	AL355098	755	15.2	76.0	175939	2	AC021191	AC021191 Homo sapi
683	15.2	76.0	153339	9	AC007350	Human chr	756	15.2	76.0	176759	2	AC037464	AC037464 Homo sapi
684	15.2	76.0	154415	2	AC108886	Rattus no	757	15.2	76.0	176798	2	AC149140	AC149140 Xenopus t
685	15.2	76.0	155540	5	BX324162	Zebrafish	758	15.2	76.0	177888	2	AC131687	AC131687 Mus muscu
686	15.2	76.0	155974	2	CRS56723	CRS56723	759	15.2	76.0	178255	9	AC019197	AC019197 Homo sapi
687	15.2	76.0	155977	10	AC112157	Danio rer	760	15.2	76.0	178272	10	AC118772	AC118772 Rattus no
688	15.2	76.0	156564	2	AC120103	Rattus no	761	15.2	76.0	178477	2	AC022801	AC022801 Homo sapi
689	15.2	76.0	156647	8	AC013790	AC137930	762	15.2	76.0	178797	2	AC121048	AC121048 Rattus no
690	15.2	76.0	157666	9	AC006989	Homo sapi	763	15.2	76.0	179526	2	HS51219	AL445624 Homo sapi
691	15.2	76.0	158231	9	AC142296	AC142296	764	15.2	76.0	180314	2	AC146158	AC146158 Pan trogl
692	15.2	76.0	158828	9	AL590069	Human DNA	765	15.2	76.0	180335	2	AC146276	AC146276 Par trogl
693	15.2	76.0	159089	2	CRS98984	Danio rer	766	15.2	76.0	181097	2	AC118205	AC118205 Mus muscu
694	15.2	76.0	159512	9	AC145986	AC145986	767	15.2	76.0	181254	2	AC147898	AC147898 Xenopus t
695	15.2	76.0	159964	2	AC125018	Mus muscu	768	15.2	76.0	181514	2	AC147450	AC147450 Otlemur t
696	15.2	76.0	160219	9	AC146277	Pan trogl	769	15.2	76.0	181535	2	AL5924203	AL5924203 Human DNA
697	15.2	76.0	160336	9	AL356276	Human DNA	770	15.2	76.0	181859	2	AC150406	AC150406 Branchios
698	15.2	76.0	160903	9	AC092812	AC092812	771	15.2	76.0	181917	2	AC133728	AC133728 Rattus no
699	15.2	76.0	160982	2	AC128945	Rattus no	772	15.2	76.0	182486	8	AC134044	AC134044 Oryza sat
700	15.2	76.0	161155	9	AP017104	AP017104	773	15.2	76.0	182666	2	AC135682	AC135682 Rattus no
701	15.2	76.0	161248	9	AC009158	AC009158	774	15.2	76.0	182859	2	ALB45341	ALB45341 Mus muscu
702	15.2	76.0	161393	10	AL645966	AL645966	775	15.2	76.0	183136	2	AC041000	AC041000 Homo sapi
703	15.2	76.0	161443	9	AC009884	AC009884	776	15.2	76.0	183396	3	AF196779	AF196779 Homo sapi
704	15.2	76.0	161569	9	AC104145	AC104145	777	15.2	76.0	184079	3	AC008258	AC008258 Drosophila
705	15.2	76.0	161590	9	AC016494	AC016494	778	15.2	76.0	184407	2	AC122767	AC122767 Mus muscu
706	15.2	76.0	161706	2	AC021194	AC021194	779	15.2	76.0	185170	5	EX005269	EX005269 Zebrafish
707	15.2	76.0	162224	9	AC012176	AC012176	780	15.2	76.0	185190	2	EX005148	EX005148 Danio rer
708	15.2	76.0	162689	2	AC117145	AC117145	781	15.2	76.0	185763	3	AF152363	AF152363 Homo sapi
709	15.2	76.0	162912	9	AC002067	AC002067	782	15.2	76.0	185800	2	AC079366	AC079366 Mus muscu
710	15.2	76.0	163353	10	AL627076	AL627076	783	15.2	76.0	185835	2	AC092007	AC092007 Bos tauru
711	15.2	76.0	163442	4	AC090445	AC090445	784	15.2	76.0	186331	9	AC006511	AC006511 Homo sapi
712	15.2	76.0	163762	4	AC090445	AC090445	785	15.2	76.0	186464	2	AC142067	AC142067 Rattus no
713	15.2	76.0	163894	2	AC150398	AC150398	786	15.2	76.0	186658	2	AP001772	AP001772 Homo sapi
714	15.2	76.0	165077	10	AC084382	AC084382	787	15.2	76.0	186798	10	AC140980	AC140980 Mus muscu
715	15.2	76.0	165240	2	AC011040	AC011040	788	15.2	76.0	186875	9	AL731540	AL731540 Human DNA
716	15.2	76.0	165603	9	EX548250	EX548250	789	15.2	76.0	186875	9	AL731540	AL731540 Human DNA
717	15.2	76.0	166686	9	AC090450	AC090450	790	15.2	76.0	187294	2	AC103917	AC103917 Drosophila
718	15.2	76.0	167015	2	EX649587	EX649587	791	15.2	76.0	187421	2	AC149882	AC149882 Xenopus t
719	15.2	76.0	167040	2	AC092194	Sus scrofa	792	15.2	76.0	187941	2	AC108858	AC108858 Mus muscu
720	15.2	76.0	167543	9	AC092554	AC092554	793	15.2	76.0	188265	9	AC146204	AC146204 Pan trogl
721	15.2	76.0	167693	2	EX640399	Danio rer	794	15.2	76.0	188380	10	AC14831	AC14831 Mus muscu
722	15.2	76.0	167755	2	EX957324	Danio rer	795	15.2	76.0	188970	2	AC115749	AC115749 Mus muscu
723	15.2	76.0	167949	9	AC087821	AC087821	796	15.2	76.0	189953	9	AL445435	AL445435 Human DNA
724	15.2	76.0	168059	3	AC010027	AC010027	797	15.2	76.0	190094	2	AC113819	AC113819 Rattus no
725	15.2	76.0	168459	9	AL772289	AL772289	798	15.2	76.0	190454	9	AC117984	AC117984 Homo sapi
726	15.2	76.0	168486	9	CNS05TBN	AL163872	799	15.2	76.0	190647	10	AL627077	AL627077 Mouse DNA
727	15.2	76.0	168560	8	AC118132	AC118132	800	15.2	76.0	191342	3	AC104140	AC104140 Drosophila
728	15.2	76.0	168575	6	AR216265	AR216265	801	15.2	76.0	191656	2	AC118574	AC118574 Lemur cat
729	15.2	76.0	168575	6	AX134178	AX134178	802	15.2	76.0	192010	10	AL732320	AL732320 Mouse DNA
730	15.2	76.0	168575	9	AY062031	AY062031	803	15.2	76.0	192429	2	AC113955	AC113955 Mus muscu
731	15.2	76.0	168764	2	AC135451	AC135451	804	15.2	76.0	192768	9	AC016831	AC016831 Homo sapi
732	15.2	76.0	168838	2	AC024929	AC024929	805	15.2	76.0	192971	10	AC132367	AC132367 Mus muscu
733	15.2	76.0	169216	2	AC090238	AC090238	806	15.2	76.0	193404	9	AC103693	AC103693 Homo sapi
734	15.2	76.0	169434	2	AC150421	AC150421	807	15.2	76.0	193511	2	AC125907	AC125907 Rattus no
735	15.2	76.0	169758	2	BX511118	BX511118	808	15.2	76.0	193514	2	AC121674	AC121674 Rattus no
736	15.2	76.0	169876	10	AL713854	AL713854	809	15.2	76.0	193813	2	AC100723	AC100723 Mus muscu
737	15.2	76.0	170085	2	AC118718	AC118718	810	15.2	76.0	194995	10	AL627103	AL627103 Mouse DNA
738	15.2	76.0	170695	2	AC068545	AC068545	811	15.2	76.0	195205	2	AC131416	AC131416 Rattus no
739	15.2	76.0	170847	2	AC144771	AC144771	812	15.2	76.0	195428	10	AC098884	AC098884 Mus muscu
740	15.2	76.0	170892	5	CNS07EES	AL445885	813	15.2	76.0	195441	2	AC091622	AC091622 Pan trogl
741	15.2	76.0	170971	5	BX572629	BX572629	814	15.2	76.0	196797	10	AC098734	AC098734 Mus muscu
742	15.2	76.0	171215	2	BX950173	BX950173	815	15.2	76.0	196852	9	AC098366	AC098366 Homo sapi
743	15.2	76.0	171612	9	AL954713	AL954713	816	15.2	76.0	197908	2	AC098366	AC098366 Rattus no
744	15.2	76.0	171734	10	AP001786	AP001786	817	15.2	76.0	198309	2	AC128119	AC128119 Rattus no
745	15.2	76.0	171744	5	EX649268	EX649268	818	15.2	76.0	198402	8	ATC87161	ATC87161 Arabidops
746	15.2	76.0	172409	5	EX649268	EX649268	819	15.2	76.0	199184	2	AC023272	AC023272 Homo sapi
747	15.2	76.0	172742	2	AC023406	AC023406	820	15.2	76.0	199240	10	AC102432	AC102432 Mus muscu
748	15.2	76.0	172962	10	AL713865	AL713865	821	15.2	76.0	199559	10	AL954362	AL954362 Mouse DNA
749	15.2	76.0	173394	9	CNS00001	AL049834	822	15.2	76.0	199923	10	AC122345	AC122345 Mus muscu

823	15.2	76.0	199946	9	AC087272	AC087272 Homo sapi	C 896	15.2	76.0	228492	10	AC132058	AC132058 Rattus no
824	15.2	76.0	200350	10	AC140208	AC140208 Mus muscu	C 897	15.2	76.0	228647	2	AC134289	AC134289 Rattus no
825	15.2	76.0	200785	2	AC108839	AC108839 Mus muscu	C 898	15.2	76.0	228825	2	AC142364	AC142364 Rattus no
826	15.2	76.0	200979	9	AC090208	AC090208 Homo sapi	C 899	15.2	76.0	230525	2	AC096982	AC096982 Rattus no
827	15.2	76.0	201412	9	AC096921	AC096921 Homo sapi	C 900	15.2	76.0	230734	2	AC115315	AC115315 Rattus no
828	15.2	76.0	201500	10	AL645807	AL645807 Mouse DNA	C 901	15.2	76.0	230738	2	AC129135	AC129135 Rattus no
829	15.2	76.0	202645	2	AL513330	AL513330 Homo sapi	C 902	15.2	76.0	230892	10	AC111721	AC111721 Mus muscu
830	15.2	76.0	202661	9	AC144386	AC144386 Pan trogl	C 903	15.2	76.0	230960	2	AC129802	AC129802 Rattus no
831	15.2	76.0	202991	10	AC149223	AC149223 Mus muscu	C 904	15.2	76.0	231542	2	AC094709	AC094709 Rattus no
832	15.2	76.0	205374	10	AC147160	AC147160 Mus muscu	C 905	15.2	76.0	231671	2	AC094705	AC094705 Rattus no
833	15.2	76.0	205621	10	AC079082	AC079082 Mus muscu	C 906	15.2	76.0	232155	2	AC111392	AC111392 Rattus no
834	15.2	76.0	205656	2	AC067966	AC067966 Homo sapi	C 907	15.2	76.0	232193	2	AC118152	AC118152 Rattus no
835	15.2	76.0	205691	2	AC087227	AC087227 Mus muscu	C 908	15.2	76.0	232473	2	AC130163	AC130163 Rattus no
836	15.2	76.0	205737	2	AC021445	AC021445 Mus muscu	C 909	15.2	76.0	233363	2	AC098461	AC098461 Rattus no
837	15.2	76.0	205972	2	AC143415	AC143415 Macaca mu	C 910	15.2	76.0	233476	10	AC107662	AC107662 Rattus no
838	15.2	76.0	206112	2	AC084412	AC084412 Mus muscu	C 911	15.2	76.0	234778	2	AC097097	AC097097 Rattus no
839	15.2	76.0	206310	9	AY371697	AY371697 Homo sapi	C 912	15.2	76.0	234804	2	AC133409	AC133409 Rattus no
840	15.2	76.0	206459	2	AC150643	AC150643 Bos tauru	C 913	15.2	76.0	234960	2	AC130922	AC130922 Rattus no
841	15.2	76.0	206748	10	AL807804	AL807804 Mouse DNA	C 914	15.2	76.0	235544	10	AC138288	AC138288 Mus muscu
842	15.2	76.0	206810	2	AC144815	AC144815 Mus muscu	C 915	15.2	76.0	235820	2	AC096613	AC096613 Rattus no
843	15.2	76.0	207534	2	AC129235	AC129235 Rattus no	C 916	15.2	76.0	236836	2	AC108530	AC108530 Rattus no
844	15.2	76.0	207901	2	AC147043	AC147043 Pan trogl	C 917	15.2	76.0	237859	2	AC124899	AC124899 Rattus no
845	15.2	76.0	208293	2	AC127052	AC127052 Rattus no	C 918	15.2	76.0	238061	2	AC109500	AC109500 Rattus no
846	15.2	76.0	208311	2	AC105301	AC105301 Mus muscu	C 919	15.2	76.0	238100	2	AC121201	AC121201 Rattus no
847	15.2	76.0	208526	2	AC149228	AC149228 Pan trogl	C 920	15.2	76.0	238521	2	AC078932	AC078932 Rattus no
848	15.2	76.0	208661	10	AC072048	AC072048 Mus muscu	C 921	15.2	76.0	238658	10	AC079818	AC079818 Mus muscu
849	15.2	76.0	208757	10	AC100752	AC100752 Mus muscu	C 922	15.2	76.0	238937	2	AC106652	AC106652 Rattus no
850	15.2	76.0	209151	2	AC092090	AC092090 Canis fam	C 923	15.2	76.0	239125	2	AC094319	AC094319 Rattus no
851	15.2	76.0	209181	2	AC127124	AC127124 Rattus no	C 924	15.2	76.0	239718	2	AC101774	AC101774 Rattus no
852	15.2	76.0	209276	2	CR388041	CR388041 Danio rer	C 925	15.2	76.0	240216	2	AC111721	AC111721 Rattus no
853	15.2	76.0	211254	2	AC129726	AC129726 Rattus no	C 926	15.2	76.0	241039	2	AC113750	AC113750 Rattus no
854	15.2	76.0	211509	9	CNS08CBD	AL928654 Human chr	C 927	15.2	76.0	242907	2	AC106660	AC106660 Rattus no
855	15.2	76.0	211546	2	AC135554	AC135554 Gallus ga	C 928	15.2	76.0	243182	2	AC120744	AC120744 Rattus no
856	15.2	76.0	211696	2	AC099580	AC099580 Mus muscu	C 929	15.2	76.0	243360	2	AC109684	AC109684 Rattus no
857	15.2	76.0	212208	2	AC079024	AC079024 Mus muscu	C 930	15.2	76.0	243768	2	AC125725	AC125725 Rattus no
858	15.2	76.0	213338	2	AC141524	AC141524 Rattus no	C 931	15.2	76.0	244105	2	AC098897	AC098897 Rattus no
859	15.2	76.0	213356	2	AC091030	AC091030 Homo sapi	C 932	15.2	76.0	244946	2	AC111788	AC111788 Mus muscu
860	15.2	76.0	213564	2	AC091030	AC091030 Homo sapi	C 933	15.2	76.0	245136	2	AC111706	AC111706 Rattus no
861	15.2	76.0	213826	10	AC113135	AC113135 Mus muscu	C 934	15.2	76.0	245174	2	AC125747	AC125747 Rattus no
862	15.2	76.0	214272	2	EX927135	EX927135 Danio rer	C 935	15.2	76.0	245794	2	AC094941	AC094941 Rattus no
863	15.2	76.0	214302	2	AC115639	AC115639 Rattus no	C 936	15.2	76.0	245965	2	AC129993	AC129993 Rattus no
864	15.2	76.0	214425	2	AC092091	AC092091 Canis fam	C 937	15.2	76.0	245966	2	AC137050	AC137050 Rattus no
865	15.2	76.0	214567	10	AC105305	AC105305 Mus muscu	C 938	15.2	76.0	246402	2	AC131535	AC131535 Rattus no
866	15.2	76.0	215697	2	AC103050	AC103050 Rattus no	C 939	15.2	76.0	246543	2	AC097034	AC097034 Rattus no
867	15.2	76.0	215876	2	AC115909	AC115909 Mus muscu	C 940	15.2	76.0	246635	2	AC094740	AC094740 Rattus no
868	15.2	76.0	216078	2	AC115165	AC115165 Rattus no	C 941	15.2	76.0	247306	2	AC131445	AC131445 Rattus no
869	15.2	76.0	216252	2	AC135266	AC135266 Rattus no	C 942	15.2	76.0	248567	2	AC128997	AC128997 Rattus no
870	15.2	76.0	216736	2	AC124080	AC124080 Mus muscu	C 943	15.2	76.0	250324	2	AC130555	AC130555 Rattus no
871	15.2	76.0	216869	2	AC137412	AC137412 Rattus no	C 944	15.2	76.0	251997	2	AC128370	AC128370 Rattus no
872	15.2	76.0	216932	10	AC090656	AC090656 Mus muscu	C 945	15.2	76.0	252129	2	AC133975	AC133975 Rattus no
873	15.2	76.0	216999	5	AC136741	AC136741 Mus muscu	C 946	15.2	76.0	252563	10	AC102091	AC102091 Mus muscu
874	15.2	76.0	217065	2	EX601644	EX601644 Zebrafish	C 947	15.2	76.0	253261	2	AC127737	AC127737 Rattus no
875	15.2	76.0	217193	2	AC110975	AC110975 Rattus no	C 948	15.2	76.0	253401	1	AC103499	AC103499 Rattus no
876	15.2	76.0	217325	10	AC115360	AC115360 Mus muscu	C 949	15.2	76.0	254050	2	AL627269	AL627269 Salmo nel
877	15.2	76.0	217801	2	AC134164	AC134164 Rattus no	C 950	15.2	76.0	254388	2	AC109427	AC109427 Rattus no
878	15.2	76.0	219121	2	AC134718	AC134718 Rattus no	C 951	15.2	76.0	254926	2	AC130105	AC130105 Rattus no
879	15.2	76.0	220678	2	AC023857	AC023857 Homo sapi	C 952	15.2	76.0	254926	2	AC101625	AC101625 Homo sapi
880	15.2	76.0	221460	2	AC126434	AC126434 Mus muscu	C 953	15.2	76.0	256320	9	AC016525	AC016525 Rattus no
881	15.2	76.0	221516	2	AC095936	AC095936 Rattus no	C 954	15.2	76.0	256643	2	AC132701	AC132701 Rattus no
882	15.2	76.0	221621	2	AC113317	AC113317 Mus muscu	C 955	15.2	76.0	256643	2	AC122567	AC122567 Rattus no
883	15.2	76.0	221722	2	AC130567	AC130567 Rattus no	C 956	15.2	76.0	259650	2	AC107176	AC107176 Rattus no
884	15.2	76.0	221899	2	AC079548	AC079548 Mus muscu	C 957	15.2	76.0	260692	2	AC142216	AC142216 Mus muscu
885	15.2	76.0	222498	2	AC094004	AC094004 Rattus no	C 958	15.2	76.0	260909	2	AC105865	AC105865 Rattus no
886	15.2	76.0	222856	2	AC128437	AC128437 Rattus no	C 959	15.2	76.0	262048	10	AC093483	AC093483 Mus muscu
887	15.2	76.0	223041	10	AC108942	AC108942 Mus muscu	C 960	15.2	76.0	262061	2	AC117172	AC117172 Rattus no
888	15.2	76.0	223850	2	AC116397	AC116397 Mus muscu	C 961	15.2	76.0	262351	2	AC105703	AC105703 Rattus no
889	15.2	76.0	223859	2	EX927287	EX927287 Danio rer	C 962	15.2	76.0	263603	2	AC128819	AC128819 Rattus no
890	15.2	76.0	224805	2	AC094882	AC094882 Rattus no	C 963	15.2	76.0	264033	1	AC133247	AC133247 Rattus no
891	15.2	76.0	224868	2	AC125933	AC125933 Rattus no	C 964	15.2	76.0	265050	2	AP000988	AP000988 Sulfolobu
892	15.2	76.0	226259	2	AC129784	AC129784 Rattus no	C 965	15.2	76.0	265099	2	AC094795	AC094795 Rattus no
893	15.2	76.0	227182	2	AC099283	AC099283 Rattus no	C 966	15.2	76.0	266673	2	AC115671	AC115671 Rattus no
894	15.2	76.0	227341	2	AC127771	AC127771 Rattus no	C 967	15.2	76.0	267054	2	AC112860	AC112860 Rattus no
895	15.2	76.0	228108	2	AC142063	AC142063 Rattus no	C 968	15.2	76.0	267182	2	AC094354	AC094354 Rattus no

```

969 15.2 76.0 267546 2 AC125740
970 15.2 76.0 267762 2 AC120693
971 15.2 76.0 269471 2 AC134107
972 15.2 76.0 270654 2 AC096084
973 15.2 76.0 270699 2 AC133821
974 15.2 76.0 271701 2 AC129343
975 15.2 76.0 272100 3 AE003840
976 15.2 76.0 272169 3 AC128071
977 15.2 76.0 274933 2 AC121671
978 15.2 76.0 278087 2 AC098278
979 15.2 76.0 281197 2 AC114571
980 15.2 76.0 283631 2 AC11536
981 15.2 76.0 284268 2 AE003565
982 15.2 76.0 285524 2 AC105802
983 15.2 76.0 287526 2 AC117308
984 15.2 76.0 290319 1 AE016982
985 15.2 76.0 295463 1 AC095332
986 15.2 76.0 295760 2 AC125718
987 15.2 76.0 296914 2 AC127750
988 15.2 76.0 297581 2 AC120937
989 15.2 76.0 298531 2 AC091363
990 15.2 76.0 298782 2 AC006844
991 15.2 76.0 300029 1 AE01553
992 15.2 76.0 300029 1 AE016839
993 15.2 76.0 300275 1 AB016751
994 15.2 76.0 300750 1 AP001509
995 15.2 76.0 309052 3 AE003425
996 15.2 76.0 312366 2 AC068766
997 15.2 76.0 322778 2 AC105550
998 15.2 76.0 326301 1 NM6622491
999 15.2 76.0 327507 2 AC096293
1000 15.2 76.0 332760 2 AC147293

```

ALIGNMENTS

```

RESULT 1 20 bp DNA linear
AX481360
LOCUS
DEFINITION Sequence 7 from Patent EP1225232.
ACCESSION AX481360
VERSION AX481360.1 GI:22316281
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Rubin, B.Y. and Anderson, S.L.
TITLE
Detection of mutations in a gene encoding Ikappab
kinase-complex-associated protein to diagnose familial dysautonomia
JOURNAL
Patent: EP 1225232-A 7 24-JUL-2002;
Rubin, Berish Y. (US); Anderson, Silvia L. (US)
LOCATION/Qualifiers
1. 20
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

PART 16-AUG-2002

ORIGIN

```

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGTCGCAACAGTACATGG 20
Db 1 AGTCGCAACAGTACATGG 20

```

```

RESULT 2
AX676048/c
LOCUS AX676048 66479 bp DNA linear PAT 27-MAR-2003

```

```

DEFINITION Sequence 1 from Patent WO02059381.
ACCESSION AX676048
VERSION AX676048.1 GI:29333739
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Slangenhuysen, S. and Gusejra, J.F.
TITLE
Gene for identifying individuals with familial dysautonomia
JOURNAL
Patent: WO 02059381-A 1 01-AUG-2002;
The General Hospital Corporation (US)
LOCATION/Qualifiers
1. 66479
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

ORIGIN

```

Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AGTCGCAACAGTACATGG 20
Db 34222 AGTCGCAACAGTACATGG 34203

```

```

RESULT 3
AL359692
LOCUS
DEFINITION Human DNA sequence from clone RP11-3J11 on chromosome 9, complete
sequence.
ACCESSION AL359692
VERSION AL359692.9 GI:14970800
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 78376)
AUTHORS
Sycamore, N.
TITLE
Direct Submission
JOURNAL
Submitted (18-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humgen@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Jul 19, 2001 this sequence version replaced gi:14715342.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-3J11 is from the library RP11-11.1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

```

http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-3311. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-3311 is at 1 in this sequence. The true left end of clone RP11-11502 is at 76377 in this sequence. The true right end of clone RP11-335N8 is at 76382 in this sequence.

FEATURES

```
source
1..78376
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-3311"
/clone_1fb="RP11-11.1"
7..829
/feature="L1PB1 repeat: matches 5343..6155 of consensus"
991..1328
/feature="MER46B repeat: matches 3..231 of consensus"
2199..2234
/feature="18 copies 2 mer tt 86% conserved"
3191..3349
/feature="AluSx/X repeat: matches 153..311 of consensus"
4250..4537
/feature="AluSx repeat: matches 1..288 of consensus"
5314..5579
/feature="AluYb repeat: matches 1..299 of consensus"
5693..5863
/feature="L1WC4 repeat: matches 6481..6679 of consensus"
5881..6013
/feature="AluYb repeat: matches 1..135 of consensus"
6533..6644
/feature="MIR repeat: matches 92..211 of consensus"
7645..7979
/feature="MIR repeat: matches 3..363 of consensus"
9165..9774
/feature="MER21B repeat: matches 179..787 of consensus"
9775..9848
/feature="MER84 repeat: matches 1..76 of consensus"
9849..10232
/feature="MER84 repeat: matches 3..373 of consensus"
10233..10535
/feature="AluSx repeat: matches 1..302 of consensus"
10536..10668
/feature="MER84 repeat: matches 373..508 of consensus"
10669..10846
/feature="MER21B repeat: matches 7..185 of consensus"
11168..11901
/feature="L1PA14 repeat: matches 5403..6149 of consensus"
12172..12468
/feature="AluSg repeat: matches 1..297 of consensus"
13074..13362
/feature="AluSx repeat: matches 3..294 of consensus"
13411..13533
/feature="MIR repeat: matches 117..239 of consensus"
13616..13669
/feature="L2 repeat: matches 2445..2498 of consensus"
13733..13841
/feature="MER53 repeat: matches 89..188 of consensus"
13996..14113
/feature="AluSg/X repeat: matches 178..295 of consensus"
14769..15142
/feature="MSTR repeat: matches 1..423 of consensus"
15274..15333
/feature="MER5A repeat: matches 14..74 of consensus"
15460..15600
/feature="MERB repeat: matches 54..175 of consensus"
15640..15689
/feature="MSTR repeat: matches 1..49 of consensus"
15881..16180
/feature="AluYb repeat: matches 1..302 of consensus"
17189..17296
```

```
repeat_region
17614..17738
/feature="MIR repeat: matches 100..206 of consensus"
17769..17834
/feature="MER5B repeat: matches 54..174 of consensus"
17769..18234
/feature="L1MD1 repeat: matches 5748..6223 of consensus"
18235..18529
/feature="AluYb repeat: matches 1..292 of consensus"
18530..19071
/feature="L1MD1 repeat: matches 5238..5748 of consensus"
19132..19432
/feature="AluSg repeat: matches 1..313 of consensus"
19441..19472
/feature="16 copies 2 mer tt 84% conserved"
19498..19859
/feature="TH1A repeat: matches 1..353 of consensus"
19860..21021
/feature="TH1A-internal repeat: matches 427..1580 of consensus"
21022..21314
/feature="AluSg repeat: matches 1..302 of consensus"
21315..21711
/feature="TH1A-internal repeat: matches 29..427 of consensus"
21718..21818
/feature="TH1B repeat: matches 265..364 of consensus"
21813..21845
/feature="Alu repeat: matches 3..35 of consensus"
21878..22184
/feature="AluSg repeat: matches 1..311 of consensus"
22578..22715
/feature="MIR repeat: matches 74..210 of consensus"
22693..22848
/feature="L2 repeat: matches 2572..2745 of consensus"
24368..24434
/feature="MIR repeat: matches 71..137 of consensus"
25259..25338
/feature="L2 repeat: matches 2605..2695 of consensus"
25474..25775
/feature="AluSg repeat: matches 1..301 of consensus"
26871..27694
/feature="CpG island"
/feature="evidence=not_experimental"
28479..28806
/feature="L2 repeat: matches 2157..2496 of consensus"
28807..29114
/feature="AluSg repeat: matches 1..305 of consensus"
29115..29296
/feature="L2 repeat: matches 1963..2157 of consensus"
29748..30050
/feature="AluSg repeat: matches 1..303 of consensus"
30475..30548
/feature="MIR repeat: matches 33..107 of consensus"
30654..30701
/feature="24 copies 2 mer ac 95% conserved"
31870..32343
/feature="L1TD repeat: matches 16..505 of consensus"
33199..33414
/feature="L2 repeat: matches 2116..2353 of consensus"
33415..33568
/feature="MER5B repeat: matches 186..340 of consensus"
33658..33918
/feature="L2 repeat: matches 2403..2710 of consensus"
37241..37362
/feature="Char11e4 repeat: matches 1823..1956 of consensus"
37669..37895
/feature="MER20 repeat: matches 1..218 of consensus"
38589..38875
/feature="AluSg repeat: matches 1..310 of consensus"
39767..39947
/feature="MSTR repeat: matches 1..183 of consensus"
40006..40255
/feature="match: STS: Em:G14836"
40477..40798
```

/note="match: STS: Em:G21606"
repeat_region 41261..41395
/note="LINE2 repeat: matches 6041..6161 of consensus"
repeat_region 41739..41805
/note="TIGER2 repeat: matches 2653..2718 of consensus"
repeat_region 41806..42093
/note="TIGER2 repeat: matches 1..300 of consensus"
repeat_region 42094..43429
/note="TIGER2 repeat: matches 1302..2653 of consensus"
repeat_region 43429..44332
/note="TIGER2 repeat: matches 297..1113 of consensus"
repeat_region 44353..44664

Query Match 100.0%; Score 20; DB 9; Length 78376;
Best Local Similarity 100.0%; Pred. No. 5.1; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTGCACAAAGTACATG 20
Db 71625 AGTGCACAAAGTACATG 71644

RESULT 4

AC148981_2/c

WPCOMMENT

Sequence split into 4 fragments LOCUS AC148981 Accession AC148981

Fragment Name	Begin	End
AC148981_0	1	110000
AC148981_1	100001	210000
AC148981_2	200001	310000
AC148981_3	300001	359936

Continuation (3 of 4) of AC148981 from base 200001 (AC148981 Mus musculus chromosome 7 c

Query Match 87.0%; Score 17.4; DB 2; Length 110000;
Best Local Similarity 94.7%; Pred. No. 1.5e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTGCACAAAGTACATG 19
Db 92061 AGTGCACAAAGTACATG 92043

RESULT 5

AC073603/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC073603 135276 bp DNA linear HTG 06-SEP-2000
Mus musculus clone RP23-6814, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
AC073603 GI:9972336
HTG; HTGS_PHASE1; HTGS_DRAFT.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 135276)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (26-JUN-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
On Sep 6, 2000 this sequence version replaced gi:9797698.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: <http://www.cshl.org/genseq>

CONTACT: mcombie@cshl.org
----- Project Information
Center project name: RP23-6814
Center clone name: RP23-6814

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 19589: contig of 19589 bp in length
* 20135: gap of unknown length
* 39580: contig of 19441 bp in length
* 39581: gap of unknown length
* 40130: gap of unknown length
* 58181: contig of 18188 bp in length
* 58319: gap of unknown length
* 58868: gap of unknown length
* 58869: contig of 16665 bp in length
* 75534: gap of unknown length
* 76083: gap of unknown length
* 92119: contig of 16036 bp in length
* 92670: gap of unknown length
* 92671: gap of unknown length
* 101790: contig of 9121 bp in length
* 102340: gap of unknown length
* 102341: contig of 6918 bp in length
* 109259: gap of unknown length
* 109808: gap of unknown length
* 116229: contig of 6421 bp in length
* 116230: gap of unknown length
* 116780: contig of 6284 bp in length
* 123064: gap of unknown length
* 123613: contig of 3929 bp in length
* 127541: contig of 3929 bp in length
* 127542: gap of unknown length
* 128090: gap of unknown length
* 128091: contig of 3264 bp in length
* 131354: gap of unknown length
* 131355: gap of unknown length
* 131904: contig of 2715 bp in length
* 134618: gap of unknown length
* 135167: gap of unknown length
* 135168: contig of 109 bp in length.
Location/Qualifiers
1..135276
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-6814"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 135276;
Best Local Similarity 94.7%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTGCACAAAGTACATG 19
Db 2982 AGTGCACAAAGTACATG 2964

RESULT 6

AC140732/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC140732 151366 bp DNA linear HTG 27-MAR-2003
Rattus norvegicus clone CH230-24B6, WORKING DRAFT SEQUENCE, 17
unordered pieces.
AC140732 GI:28626532
HTG; HTGS_PHASE1; HTGS_DRAFT.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 151366)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Murny, D., Marie, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alstbrooks, S., Amin, A., Arguiano, D.,
Ayalebech, V., Ayagi, A., Ayodeji, I., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacco, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, K., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hernandez, U., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, U., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louised, H., Lozada, R.J., Lu, X., Ma, V., Maheswari, M., Mahindaratne, N., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munday, M., Norris, S., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokeneme, O., Okunori, G., Olarinmwa, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sma, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, D., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, R., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, K., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 151366)
Worley, K.C.
Direct Submission
Submitted (01-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 151366)
Worley, K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEBF
Center clone name: CH230-24B6
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phred; version 0.990329
Consensus quality: 143638 bases at least Q40
Consensus quality: 145076 bases at least Q30
Consensus quality: 145959 bases at least Q20

Estimated insert size: 143140; sum-of-coverage estimation
Quality coverage: 4x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1151: contig of 1151 bp in length
* 1152 1251: gap of unknown length
* 1152 2326: contig of 1075 bp in length
* 2327 2426: gap of unknown length
* 2427 3587: contig of 1161 bp in length
* 3588 3687: gap of unknown length
* 3688 5496: contig of 1809 bp in length
* 5497 5597: gap of unknown length
* 5597 8627: contig of 3031 bp in length
* 8628 8727: gap of unknown length
* 8728 13024: contig of 4297 bp in length
* 13025 13124: gap of unknown length
* 13125 20686: contig of 7562 bp in length
* 20687 20786: gap of unknown length
* 20787 28708: contig of 7922 bp in length
* 28709 28809: gap of unknown length
* 28809 37018: contig of 8210 bp in length
* 37019 37118: gap of unknown length
* 37119 45173: contig of 8055 bp in length
* 45174 45273: gap of unknown length
* 45274 56516: contig of 11243 bp in length
* 56517 56616: gap of unknown length
* 56617 69433: contig of 12817 bp in length
* 69434 79537: gap of unknown length
* 79538 79557: contig of 10024 bp in length
* 79558 79657: gap of unknown length
* 79658 96744: contig of 17087 bp in length
* 96745 96745: gap of unknown length
* 96845 110551: contig of 13707 bp in length
* 110552 110651: gap of unknown length
* 110652 128099: contig of 17448 bp in length
* 128100 128199: gap of unknown length
* 128200 151366: contig of 23167 bp in length.
*
FEATURES
source
1. 151366
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-24B6"
ORIGIN
Query Match 87.0% Score 17.4; DB 2; Length 151366;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GTGGCAACAGTACATGG 20
DB 103450 GTGGCAACAGTACATGG 103432
RESULT 7
AC099737 202183 bp DNA linear HTG 03-MAR-2004
LOCUS Mus musculus chromosome 17 clone RP23-247M16 map 17, *** SEQUENCING
DEFINITION IN PROGRESS ***, 8 unordered pieces.
ACCESSION AC099737.3 GI:44890998
VERSION AC099737
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULFIP; HTGS_ACTIVEFIN.
SOURCE
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 202183)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 17, clone RP23-247M16

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 202183)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campio, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Grande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Jones, C., Kamat, A., Karatas, A., Kellis, C., Larocque, K., Lamas, R., Landers, T., Lebeck, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (19-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 202183)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Atacheni, H., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rieback, M., Retta, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (03-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2004 this sequence version replaced gr1:28475922.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L17629
Center clone name: 247_M_16

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 39687: contig of 39687 bp in length
* 39688 39787: gap of 100 bp
* 39788 80800: contig of 41013 bp in length
* 80801 80900: gap of 100 bp
* 80901 131062: contig of 50162 bp in length
* 131063 140256: contig of 9094 bp in length
* 140257 140356: gap of 100 bp
* 140357 146997: contig of 6641 bp in length
* 146998 147097: gap of 100 bp
* 147098 172292: contig of 25195 bp in length
* 172293 173392: gap of 100 bp
* 173393 191398: contig of 19006 bp in length
* 191399 191498: gap of 100 bp
* 191499 202183: contig of 10685 bp in length.
Location/Qualifiers
1..202183
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP23-247M16"
/clone_lib="RP23-247M16" Female Mouse BAC"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 202183;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTGGCAACGCTACATG 19
Db 128238 AGTCCAAACGCTACATG 128256

RESULT 8
AC119471/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-414D17, WORKING DRAFT SEQUENCE, 5
nucleated pieces.
AC119471
AC119471.5 GI:25098850
VERSION
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 205681)
Muzny, D., Marle, Mezker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D., Ayralbechi, V., Ayragi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blatt, J., Blankenburg, K., Blayth, P., Brown, M., Bryant, N., Buhay, C., Burck, P., Burrelli, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyte, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenhubwa, L., Loulsegod, H., Lozada, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
 Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Miosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okunnu, G., Olarinmuse, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C.,
 Pioppert, F., Polndexter, A., Popovic, D., Primus, E., Pu, L.,
 Piazzi, M., Quito, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R.,
 Ralby, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 205681)
 Worley, K.C.
 Direct Submission
 Submitted (27-Apr-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 205681)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-Nov-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:2390845.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GVEP
 Center clone name: CH230-414D17
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 192343 bases at least Q40
 Consensus quality: 19463 bases at least Q30
 Consensus quality: 195894 bases at least Q20
 Estimated insert size: 198896; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 196852: contig of 196852 bp in length
 * 196853 196952: gap of unknown length
 * 196953 198215: contig of 1263 bp in length
 * 198216 198315: gap of unknown length
 * 198316 200209: contig of 1894 bp in length
 * 200210 200309: gap of unknown length
 * 200310 202037: contig of 1728 bp in length
 * 202038 202137: gap of unknown length
 * 202138 205681: contig of 3544 bp in length.

FEATURES
 source
 1. 205681
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-414D17"
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77
 site:"
 end_sequence:B2170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:"
 end_sequence:B2170609"

misc_feature
 misc_feature
 misc_feature
 misc_feature
 misc_feature

ORIGIN
 Query Match 87.0%; Score 17.4; DB 2; Length 205681;
 Best Local Similarity 94.7%; Pred. No. 1.4e+02;
 Matches 18; Conservative 0; Mismatches 1; Gaps 0;

QY 2 GTCCGAACAGTCAATG 20
 Db 70462 GTCCGACAGTCAATG 70444

RESULT 9
 AC073796 215196 bp DNA linear HTG 29-JUN-2000
 LOCUS Mus musculus clone R23-431K23, WORKING DRAFT SEQUENCE, 27
 DEFINITION Mus musculus clone R23-431K23, WORKING DRAFT SEQUENCE, 27
 ACCESSION AC073796
 VERSION AC073796.1 GI:8810413
 KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCES
 1 (bases 1 to 215196)
 DOE Joint Genome Institute.
 JOURNAL Sequencing of Mouse
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 215196)
 TITLE Direct Submission
 JOURNAL DOE Joint Genome Institute.
 AUTHORS Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1902639
Center clone name: RPCR-23_431K23

Summary Statistics

Consensus quality: 193161 bases at least Q40
Consensus quality: 201982 bases at least Q30
Consensus quality: 204342 bases at least Q20
Estimated insert size: 190000; agarose-gel estimation
Estimated insert size: 212596; sum-of-contrigs estimation
Quality coverage: 8.52 in Q20 bases; agarose-gel estimation
Quality coverage: 7.62 in Q20 bases; sum-of-contrigs estimation.

NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1535: contrig of 1535 bp in length
1536 1635: gap of unknown length
1636 2881: contrig of 1246 bp in length
2882 2981: gap of unknown length
2982 4237: contrig of 1256 bp in length
4238 4338: gap of unknown length
4338 5459: contrig of 1122 bp in length
5459 5559: gap of unknown length
5560 6642: contrig of 1083 bp in length
6643 6743: gap of unknown length
6743 8140: contrig of 1398 bp in length
8141 8240: gap of unknown length
8241 10071: contrig of 1831 bp in length
10072 10171: gap of unknown length
10172 11363: contrig of 1192 bp in length
11364 11463: gap of unknown length
11464 14217: contrig of 2754 bp in length
14218 14317: gap of unknown length
14318 17815: contrig of 3498 bp in length
17816 17915: gap of unknown length
17916 20634: contrig of 2719 bp in length
20635 20734: gap of unknown length
20735 24318: contrig of 3584 bp in length
24319 24418: gap of unknown length
24419 27634: contrig of 3216 bp in length
27635 27734: gap of unknown length
27735 32614: contrig of 4880 bp in length
32615 32714: gap of unknown length
32715 37575: contrig of 4865 bp in length
37576 37675: gap of unknown length
37676 44946: contrig of 7267 bp in length
44947 45046: gap of unknown length
45047 50203: contrig of 5157 bp in length
50204 50303: gap of unknown length
50304 54763: contrig of 4460 bp in length
54764 54863: gap of unknown length
54864 65724: contrig of 10861 bp in length
65725 65824: gap of unknown length
65825 75736: contrig of 9912 bp in length
75737 75836: gap of unknown length
75837 84975: contrig of 9139 bp in length
84976 85075: gap of unknown length
85076 101226: contrig of 16151 bp in length
101227 101326: gap of unknown length
101327 120390: contrig of 19064 bp in length
120391 120490: gap of unknown length
120491 135635: contrig of 15145 bp in length
135636 135735: gap of unknown length
135736 162383: contrig of 26648 bp in length
162384 162483: gap of unknown length
162484 188312: contrig of 25829 bp in length
188313 188412: gap of unknown length

FEATURES

* 188413 215196: contrig of 26784 bp in length.
Location/Qualifiers
1..215196
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-431K23"
/clone_11b="RPCR mouse BAC library 23"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 215196;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservativity 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 AGTGCACACGATGACATG 19
|||||
DB 213819 AGTCCACACGATGATG 213837

RESULT 10

AC149052 215199 bp DNA linear HTG 04-AUG-2004
LOCUS Mus musculus chromosome 7 clone RP24-54801, WORKING DRAFT SEQUENCE,
DEFINITION 8 unordered pieces.
ACCESSION AC149052
VERSION AC149052.2 GI:50950386
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE

1 Wilson, R.K.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 215199)
Wilson, R.K.
Submitted (08-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 215199)
Wilson, R.K.
Direct Submission
Submitted (04-AUG-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Aug 4, 2004 this sequence version replaced gi:47084635.

REFERENCE

TITLE
JOURNAL
AUTHORS
JOURNAL
TITLE
JOURNAL

REFERENCE

TITLE
JOURNAL
AUTHORS
JOURNAL
TITLE
JOURNAL

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu

Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BB0548001

Summary Statistics

Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210989 bases at least Q40
Consensus quality: 211541 bases at least Q30
Consensus quality: 211780 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*
* 1 1024: contig of 1024 bp in length
* 1025 1124: gap of unknown length
* 1125 2558: contig of 1434 bp in length
* 2559 2658: gap of unknown length
* 2659 3853: contig of 1194 bp in length
* 3853 3953: gap of unknown length
* 3953 23853: contig of 19901 bp in length
* 23853 23954: gap of unknown length
* 23954 46296: contig of 22343 bp in length
* 46297 46397: gap of unknown length
* 46397 96500: contig of 50104 bp in length
* 96501 152173: gap of unknown length
* 152173 152272: gap of unknown length
* 152273 215199: contig of 62927 bp in length.
*
* Location/Qualifiers
* 1. 215199
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /chromosome="7"
* /clone="RP24-54801"
* 1. 1024
* /note="assembly_name:Contig12"
* 1125. 2558
* /note="assembly_name:Contig16"
* 2559. 3852
* /note="assembly_name:Contig27"
* 3953. 123853
* /note="assembly_name:Contig36"
* 23954. 46296
* /note="assembly_name:Contig37"
* clone_end:17
* vector_side:right"
* 46397. 96500
* /note="assembly_name:Contig38"
* 96501. 152172
* /note="assembly_name:Contig39"
* 152273. 215199
* /note="assembly_name:Contig40"

ORIGIN
Query Match 87.0% Score 17.4; DB 2; Length 215199;
Best Local Similarity 94.7% Pired. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCGCAACGATACATG 19
Db 113302 AGTCGCAACGATACATG 113320

RESULT 11
AC140735 229472 bp DNA linear HTG 27-MAR-2003
LOCUS Rattus norvegicus clone CH230-199G5, WORKING DRAFT SEQUENCE, 50
DEFINITION Unordered pieces.
ACCESSION AC140735
VERSION AC140735.4 GI:28894512
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229472)
Muzny,D,Marler,M,Mezker,M,Lee,S,Abramson,S,Adams,C,Alder,J,
Allen,C,Allen,H,Alsbrook,S,Amin,A,Angiano,D,
Aryalambuchi,V,Aoyagi,A,Ayodeji,M,Baca,E,Baden,H,
Baldwin,D,Bandaranaike,D,Barber,M,Barnstead,M,Benahmed,F,
Blawie,K,Blair,J,Blankenburg,K,Blyth,P,Brown,M,
Bryant,N,Buhay,C,Burch,P,Burrelli,K,Calderson,E,
Cardenas,V,Carter,K,Cavazos,I,Ceasar,H,Center,A,
Chacko,J,Chavez,D,Chen,G,Chen,R,Chen,Y,Chen,Z,Chu,J,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,B, Davis,C, Davy-Carrill,L, De Anda,C, Dedrich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Diaper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Gant,R, Garcia,A, Garner,T, Garza,M,
Georgievski,E, Geer,R, Gill,R, Grady,M, Guerra,M, Guevara,W,
Gunaratne,P, Haaland,M, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howell,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kivar,C,
Kowls,C, Kraft,C,L, Lebow,H, Levan,D, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenschewa,L, Louissege,H, Lozada,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindaratne,M, Mahmud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Manguney,S, McLeod,M, McNeill,T, Meenen,E, Mlcosavljevic,A,
Miner,G, Minja,E, Montemayor,U, Moore,S, Morgan,M, Morris,K,
Morris,S, Mundasa,M, Murphy,M, Murphy,M, Nankervis,C, Neal,D,
Newton,N, Nguyen,N, Norris,S, Nwokeneme,O, Okwono,G,
Olatunmbunsgoon,A, Pal,S, Parks,K, Pasternak,S, Paul,H,
Perez,A, Perez,L, Pfannkuch,C, Pioppey,F, Poindecker,A,
Popovic,D, Primus,B, Pu,L-L, Puzo,M, Quiroz,J, Rachlin,E,
Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y,
Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A,
Rose,M, Rose,R, Ruiz,S, Sanders,W, Severy,G, Scherer,S,
Scott,G, Shatsman,S, Shen,H, Shelly,J, Shvartsbeyn,A,
Sisson,I, Sitter,C,D, Smajs,D, Speed,A, Sodergren,E,
Song,X-Z, Sorelle,R, Sosa,J, Steime,M, Strong,R, Sutton,A,
Svatek,A, Tabors,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S,
Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,Y, Villaseana,D,
Waldron,L, Walker,B, Wang,U, Wang,Q, Wang,S, Warren,J,
Warren,R, Wei,X, White,F, Williams,G, Wilson,R, Wiegck,R,
Wooten,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S,
Yen,J, Yeon,L, Yeon,V, Yu,F, Zhang,J, Zhou,Z, Zhou,X,
Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R,
Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.

Direct Submission
Unpublished
2 (bases 1 to 229472)
Worley,K,C.
Submitted (01-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229472)
Worley,K,C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 10, 2003 this sequence version replaced gi:28894512.

----- genome center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEBB
Center clone name: CH230-199G5
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 213297 bases at least Q40
Consensus quality: 218310 bases at least Q30
Consensus quality: 22517 bases at least Q20
Estimated insert size: 211997; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

```

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 50 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1171: contig of 1171 bp in length
 1172 1271: gap of unknown length
 1272 2336: contig of 1065 bp in length
 2337 2436: gap of unknown length
 2437 3503: contig of 1067 bp in length
 3504 3603: gap of unknown length
 3604 4627: contig of 1024 bp in length
 4628 4727: gap of unknown length
 4728 5961: contig of 1234 bp in length
 5962 6061: gap of unknown length
 6062 7167: contig of 1106 bp in length
 7168 7267: gap of unknown length
 7268 8763: contig of 1496 bp in length
 8764 10004: gap of unknown length
 10005 10104: contig of 1141 bp in length
 10105 11619: contig of 1515 bp in length
 11620 11719: gap of unknown length
 11720 13508: contig of 1789 bp in length
 13509 13608: gap of unknown length
 13609 15188: contig of 1580 bp in length
 15189 15288: gap of unknown length
 15289 16965: contig of 1677 bp in length
 16966 17065: gap of unknown length
 17066 18268: contig of 1203 bp in length
 18269 20327: contig of 1959 bp in length
 20328 20427: gap of unknown length
 20428 22313: contig of 1886 bp in length
 22314 22413: gap of unknown length
 22414 23580: contig of 1167 bp in length
 23581 23680: gap of unknown length
 23681 25026: contig of 1346 bp in length
 25027 25126: gap of unknown length
 25127 26809: contig of 1683 bp in length
 26810 26909: gap of unknown length
 26910 28612: contig of 1703 bp in length
 28613 28712: gap of unknown length
 28713 30395: contig of 1683 bp in length
 30396 30495: gap of unknown length
 30496 31872: contig of 1377 bp in length
 31873 31972: gap of unknown length
 31973 33259: contig of 1287 bp in length
 33260 33359: gap of unknown length
 33360 35432: contig of 2073 bp in length
 35433 35532: gap of unknown length
 35533 36968: contig of 1436 bp in length
 36969 37068: gap of unknown length
 37069 38344: contig of 1276 bp in length
 38345 38444: gap of unknown length
 38445 40111: contig of 1567 bp in length
 40112 40112: gap of unknown length
 40113 41147: contig of 1036 bp in length
 41148 41247: gap of unknown length
 41249 43338: contig of 2091 bp in length
 43339 43438: gap of unknown length
 43439 45443: contig of 2005 bp in length
 45444 45543: gap of unknown length
 45544 48112: contig of 2565 bp in length
 48113 48212: gap of unknown length
 48213 50595: contig of 2383 bp in length
 50596 52465: gap of unknown length
 52466: contig of 1770 bp in length

52466 52565: gap of unknown length
 52566 55543: contig of 2978 bp in length
 55544 55643: gap of unknown length
 55644 58521: contig of 2878 bp in length
 58522 58621: gap of unknown length
 58622 61051: contig of 2429 bp in length
 61052 61151: gap of unknown length
 61152 64258: contig of 3108 bp in length
 64259 64358: gap of unknown length
 64359 68529: contig of 4171 bp in length
 68530 68629: gap of unknown length
 68630 72708: contig of 4079 bp in length
 72709 72809: gap of unknown length
 72810 79838: contig of 7031 bp in length
 79839 79938: gap of unknown length
 79939 79940: contig of 7074 bp in length
 79941 87113: gap of unknown length
 87114 87114: contig of 6009 bp in length
 87115 93123: gap of unknown length
 93124 93223: contig of 7290 bp in length
 93224 100512: contig of 7290 bp in length
 100513 100612: gap of unknown length
 100613 109171: contig of 8553 bp in length
 109172 109172: gap of unknown length
 109173 118889: contig of 9618 bp in length
 118890 118890: gap of unknown length
 118891 132052: contig of 13063 bp in length
 132053 132152: gap of unknown length
 132153 142708: contig of 10557 bp in length
 142709 142809: gap of unknown length
 142810 160095: contig of 17286 bp in length
 160096 160195: gap of unknown length
 160196 177073: contig of 16878 bp in length
 177074 177173: gap of unknown length
 177174 199671: contig of 22498 bp in length

Query Match 87.0%; Score 17.4; DB 2; Length 229472;
 Best Local Similarity 94.7%; Pred. No. 1,4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGGCAACAGTACATG 20
 Db 18905 GTGGCAACAGTACATG 18923

RESULT 12
 AC136831/c 243210 bp DNA linear HTG 23-NOV-2002
 LOCUS Rattus norvegicus clone CH230-7418, WORKING DRAFT SEQUENCE, 3
 DEFINITION Rattus norvegicus clone CH230-7418, WORKING DRAFT SEQUENCE, 3
 ACCESSION AC136831.2 GI:25188459
 VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus

REFERENCE
 AUTHORS
 1 (bases 1 to 243210)
 Mueny D,Marie, Mezker,M,Lee, Abramson,S, Adams,C, Alder,J,
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
 Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
 Biswal,D, Blair,U, Blankenburg,K, Blyth,P, Brown,M,
 Bryant,N, Buhay,C, Burch,P, Surrrell,K, Calderon,E,
 Cardenas,V, Carter,K, Cavazos,I, Cessari,H, Center,A,
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
 Cleveland,C, Cockrell,R, Cox,C, Coyne,M, Cree,A, D'Souza,L,
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedertich,D,
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, DiVya,K,
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
 Egan,A, Escotto,K, Eugene,C, Evans,C,A, Falls,T, Fan,G,
 Fernandez,S, Finley,M, Flegg,N, Forbes,L, Foster,M, Foster,P,
 Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,

Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulky, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, S., Johnson, R., Joliver, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lounsbury, L., Louised, R., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nsookeleneh, O., Okunolu, G., Olarunpunsagoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pr, L., L., Piaz, M., Quiroz, J., Rachin, E., Reeves, K., Register, M.A., Reigh, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojase, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J., Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, J., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 243210)
Rat Genome Sequencing Consortium.
Submitted (08-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 243210)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:24796705.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a contig-scaffold). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTO1
Center clone name: CH230-7418
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 230829 bases at least Q40
Consensus quality: 233240 bases at least Q30

Consensus quality: 234747 bases at least Q20
Estimated insert size: 236645, sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 240139: contig of 240139 bp in length
* 240140 240239: gap of unknown length
* 240240 241508: contig of 1269 bp in length
* 241509 241608: gap of unknown length
* 241609 243210: contig of 1602 bp in length.
Location/Qualifiers
1. 243210
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-7418"
1. 1466
/note="wgs_contig"

ORIGIN
misc_feature
Query Match 87.0%; Score 17.4; DB 2; Length 243210;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;

DB 74206 AGACCGAACAAGTACATG 74188

OR 1 AGTCGAAACAGTACATG 19
|||||
|||||

RESULT 13
AC115331
LOCUS
DEFINITION
Rattus norvegicus clone CH230-307D19. *** SEQUENCING IN PROGRESS
** 7 unordered pieces.
AC115331
AC115331.3 GI:23270141
HTG: HTGS_PHASE1: HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus
Rattus norvegicus
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 243702)
Muzny, D.Marie, Metzker, M.Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Altschrocks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnham, F., Bissalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Flinkey, M., Flagg, N., Forbes, J., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, J., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulky, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, S., Johnson, R., Joliver, A.,

Karpachy, S., Kelly, S., Khan, Z., King, L., Kovari, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Mallory, K., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., McManley, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mijosevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nsoekeleneh, O., Okunolu, G., Olarnunsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, L., Perez, L., Pfankuch, C., Ploger, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, P., Quirroz, J., Rachlin, E., Reeves, K., Register, M.A., Reish, R., Rella, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soga, J., Steinle, M., Strong, R., Sutton, A., Syatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemari, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 243702)
 AUTHORS Morley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 243702)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (23-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 23, 2002 this sequence version replaced gi:21781813.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Project name: G00A
 Center Project name: CH230-307D19
 Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 175747 bases at least Q40
 Consensus quality: 178942 bases at least Q30
 Consensus quality: 180494 bases at least Q20
 Estimated insert size: 199396; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 3679: contig of 3679 bp in length
 * 3680 3779: gap of unknown length
 * 3780 208206: contig of 204427 bp in length
 * 208305: gap of unknown length
 * 208307 234315: contig of 26009 bp in length
 * 234316 234415: gap of unknown length
 * 234416 235574: contig of 1159 bp in length
 * 235575 23674: gap of unknown length
 * 23675 237230: contig of 1556 bp in length
 * 237231 237330: gap of unknown length
 * 237331 238464: contig of 2134 bp in length
 * 238465 239564: gap of unknown length
 * 239565 243702: contig of 4138 bp in length.

Location/Qualifiers
 1..243702
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-307D19"
 1..1338
 /note="wgs_contig"
 3780..4825
 /note="wgs_contig"
 206630..208206
 /note="wgs_contig"

ORIGIN
 Query Match 87.0%; Score 17.4; DB 2; Length 243702;
 Best Local Similarity 94.7%; Pred. No. 1,4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCCAACAGTACATCG 20
 Db 223467 GTGGCAACAGTACATCG 223485

RESULT 14
 AC094556/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-3P4, WORKING DRAFT SEQUENCE, 5
 UNORDERED
 AC094556
 AC094556.6 GI:30466689
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 289856)
 Muray, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Atsbrooke, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chert, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L.,
 Davis, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Detam, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,

```

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, D.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, L., Lougheed, H., Lozano, R.J., Lu, X., Ma, J.,
Manshwar, M., Mandarte, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwankwelu, O., Okwum, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Piamkoc, C.,
Plapper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Piazo, M., Quito, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Rally, B., Rally, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shwartsbeyn, A., Sisson, T., Sitter, C.D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K.,
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhu, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G., and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
Worley, K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 289856)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:2492102.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAWR
Center clone name: CH230-3P4
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 238163 bases at least Q40
Consensus quality: 238828 bases at least Q30
Consensus quality: 241498 bases at least Q20
Estimated insert size: 246461; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

```

```

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 282464: contig of 282464 bp in length
* 282465 282564: gap of unknown length
* 282565 283807: contig of 1243 bp in length
* 283808 283907: gap of unknown length
* 283908 286835: contig of 2928 bp in length
* 286836 286935: gap of unknown length
* 286936 288319: contig of 1384 bp in length
* 288320 288419: gap of unknown length
* 288420 289856: contig of 1437 bp in length.

FEATURES
source
1. 289856
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3P4"
misc_feature
1..2521
/note="wgs contig"
misc_feature
108311..110551
/note="wgs contig"
ORIGIN
Query Match 87.0%; Score 17.4; DB 2; Length 289856;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCGAACAGTACATCG 20
Db 95008 GTGGCAACAGTACATCG 94990

RESULT 15
HHV6U1102 10601 bp DNA linear VRL 27-OCT-1995
LOCUS Human herpesvirus type 6A tp/cap, mdbp, U42orf, U43orf, U44orf,
DEFINITION and dUTPase genes.
ACCESSION X92436.1 GI:1044869
VERSION dUTPase; mdbp gene; tp/cap gene; U40; U41; U42; U43; U44; U45.
KEYWORDS Human herpesvirus 6
SOURCE Human herpesvirus 6
ORGANISM
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Roseolovirus.

REFERENCE
1 Jones, M.D.
TITLE DNA sequence analysis of the origin of DNA replication region of
AUTHORS Human Herpesvirus 6A strain U1102
JOURNAL Unpublished
2 (bases 1 to 10601)
REFERENCE Jones, M.D.
TITLE Direct Submission
AUTHORS Submitted (18-OCT-1995) M.D. Jones, Royal Postgraduate Medical
JOURNAL School, Virology, Du Cane Road, London, W12 0NN, UK
Location/Qualifiers
1..10601
/organism="Human herpesvirus 6"
/mol_type="genomic DNA"
/strain="U1102"
/isolate="U1102"
/db_xref="taxon:10368"
/cell_line="T cell"
/lab_host="human T-cells"
complement(1..278)
/gene="tp/cap"
complement(1..278)
CDS

```


TITLE Homo sapiens chromosome 17, clone RP11-88B16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 56984)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguski, L., Bouckgeater, B.,
Brown, A., Camarata, J., Campoliano, A., Chang, J., Charat, B.,
Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooper, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Hardford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karacas, A., Kells, C., Laroque, K.,
Lamares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McSwan, P., McKernan, K., McPherson, R., Melidim, D.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Polara, V.,
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strass, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Sult, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Project name: 122861
Center project name: 88_B_16
Center clone name: 88_B_16

* NOTE: This record contains 69 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 753: contig of 753 bp in length
* 754 853: gap of 100 bp
* 854 1596: contig of 743 bp in length
* 1597 1696: gap of 100 bp
* 1697 2438: contig of 742 bp in length
* 2439 2538: gap of 100 bp
* 2539 3278: contig of 740 bp in length
* 3279 3378: gap of 100 bp
* 3379 4106: contig of 728 bp in length
* 4107 4206: gap of 100 bp
* 4207 4890: contig of 684 bp in length
* 4891 4990: gap of 100 bp
* 4991 5701: contig of 711 bp in length
* 5702 5801: gap of 100 bp
* 5802 6549: contig of 748 bp in length
* 6550 6649: gap of 100 bp
* 6650 7353: contig of 704 bp in length
* 7354 7453: gap of 100 bp
* 7454 8144: contig of 691 bp in length
* 8145 8244: gap of 100 bp
* 8245 8991: contig of 747 bp in length
* 8992 9091: gap of 100 bp

9092 9838: contig of 747 bp in length
* 9839 9938: gap of 100 bp
* 9939 10682: contig of 744 bp in length
* 10683 10782: gap of 100 bp
* 10783 11521: contig of 739 bp in length
* 11522 12345: contig of 724 bp in length
* 12346 12446: gap of 100 bp
* 12446 13170: contig of 725 bp in length
* 13171 13270: gap of 100 bp
* 13271 13983: contig of 713 bp in length
* 13984 14083: gap of 100 bp
* 14084 14787: contig of 704 bp in length
* 14788 14887: gap of 100 bp
* 14888 15630: contig of 743 bp in length
* 15631 15730: gap of 100 bp
* 15731 16475: contig of 745 bp in length
* 16476 16575: gap of 100 bp
* 16576 17301: contig of 726 bp in length
* 17302 17401: gap of 100 bp
* 17402 18124: contig of 723 bp in length
* 18125 18224: gap of 100 bp
* 18225 18937: contig of 713 bp in length
* 18938 19037: gap of 100 bp
* 19038 19778: contig of 741 bp in length
* 19779 19878: gap of 100 bp
* 19879 20589: contig of 711 bp in length
* 20590 20689: gap of 100 bp
* 20690 21418: contig of 729 bp in length
* 21419 21518: gap of 100 bp
* 21519 22320: contig of 702 bp in length
* 22321 22320: gap of 100 bp
* 22321 23018: contig of 698 bp in length
* 23019 23118: gap of 100 bp
* 23119 23806: contig of 688 bp in length
* 23807 23906: gap of 100 bp
* 23907 24649: contig of 743 bp in length
* 24650 24749: gap of 100 bp
* 24750 25479: contig of 730 bp in length
* 25480 25579: gap of 100 bp
* 25580 26111: contig of 732 bp in length
* 26112 26411: gap of 100 bp
* 26412 27138: contig of 727 bp in length
* 27139 27238: gap of 100 bp
* 27239 27944: contig of 706 bp in length
* 27945 28044: gap of 100 bp
* 28045 28755: contig of 711 bp in length
* 28756 28855: gap of 100 bp
* 28856 29587: contig of 732 bp in length
* 29588 29687: gap of 100 bp
* 29688 30398: contig of 711 bp in length
* 30399 30498: gap of 100 bp
* 30499 31189: contig of 691 bp in length
* 31190 31289: gap of 100 bp
* 31290 32188: contig of 899 bp in length
* 32189 32288: gap of 100 bp
* 32289 33022: contig of 734 bp in length
* 33023 33122: gap of 100 bp
* 33123 33854: contig of 732 bp in length
* 33855 33954: gap of 100 bp
* 33955 34682: contig of 728 bp in length
* 34683 34782: gap of 100 bp
* 34783 35516: contig of 724 bp in length
* 35517 35616: gap of 100 bp
* 35617 36329: contig of 713 bp in length
* 36330 36429: gap of 100 bp
* 36430 37147: contig of 718 bp in length
* 37148 37247: gap of 100 bp
* 37248 37976: contig of 729 bp in length
* 37977 38076: gap of 100 bp
* 38077 38780: contig of 704 bp in length
* 38781 38880: gap of 100 bp
* 38881 39572: contig of 692 bp in length

```

* 39573 39672: gap of 100 bp
* 39673 40405: contig of 737 bp in length
* 40405 40509: gap of 100 bp
* 40509 41246: contig of 737 bp in length
* 41246 41347: gap of 100 bp
* 41347 42081: contig of 735 bp in length
* 42081 42181: gap of 100 bp
* 42181 42892: contig of 711 bp in length
* 42892 42992: gap of 100 bp
* 42992 43701: contig of 709 bp in length
* 43701 43801: gap of 100 bp
* 43801 44500: contig of 699 bp in length
* 44500 44600: gap of 100 bp
* 44600 45305: contig of 705 bp in length
* 45305 45406: gap of 100 bp
* 45406 46108: contig of 703 bp in length
* 46108 46208: gap of 100 bp
* 46208 46906: contig of 698 bp in length
* 46906 47007: gap of 100 bp
* 47007 47747: contig of 741 bp in length
* 47747 47847: gap of 100 bp
* 47847 48588: contig of 741 bp in length
* 48588 48689: gap of 100 bp
* 48689 49417: contig of 729 bp in length
* 49417 49517: gap of 100 bp
* 49517 50254: contig of 737 bp in length
* 50254 50354: gap of 100 bp
* 50354 51070: contig of 715 bp in length
* 51070 51920: contig of 751 bp in length
* 51920 52020: gap of 100 bp
* 52020 52741: contig of 721 bp in length
* 52741 52841: gap of 100 bp
* 52841 53540: contig of 698 bp in length
* 53540 53639: gap of 100 bp
* 53639 54391: contig of 752 bp in length
* 54391 54492: gap of 100 bp
* 54492 55233: contig of 742 bp in length
* 55233 55334: gap of 100 bp
* 55334 56062: contig of 729 bp in length
* 56062 56163: gap of 100 bp
* 56163 56884: contig of 722 bp in length.

```

Query Match 84.0%; Score 16.8; DB 2; Length 56884;
 Best Local Similarity 90.0%; Pred. No. 3.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGTCGCAACGATCAATGG 20
 Db 33562 AGTCGCAACGATCAAGGG 33581

RESULT 17
 AC104586
 LOCUS
 DEFINITION Homo sapiens chromosome 17 clone RP11-77P9 map 17, LOW-PHASE
 SEQUENCE SAMPLING.

AC104586 71049 bp DNA linear HTG 13-DEC-2001
 AC104586.1 GI:17646986
 VERSION HTG; HTGS_PHASE0.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 1 (bases 1 to 71049)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,

TITLE JOURNAL COMMENT

Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gargyala, S.,
 Girde, S., Gord, S., Goyette, M., Gratum, L., Grand-Pierre, N.,
 Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kanat, A., Karatas, A., Kellis, C., Labèque, R.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McMan, P., McKernan, K., McPheters, N., Meldrum, J.,
 Menene, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
 Zainoun, U., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

Project Information
 Center project name: L22858
 Center clone name: 77_P_9

NOT: This record contains 84 individual
 sequencing reads that have not been assembled into
 contigs. Runs of N are used to separate the reads
 and the order in which they appear is completely
 arbitrary. Low-pass sequence sampling is useful for
 identifying clones that may be gene-rich and allows
 overlap relationships among clones to be deduced.
 However, it should not be assumed that this clone
 will be sequenced to completion. In the event that
 the record is updated, the accession number will
 be preserved.

```

* 1 723: contig of 723 bp in length
* 724 823: gap of 100 bp
* 824 1602: contig of 779 bp in length
* 1603 1702: gap of 100 bp
* 1703 2425: contig of 723 bp in length
* 2426 2525: gap of 100 bp
* 2526 3252: contig of 727 bp in length
* 3253 3353: gap of 100 bp
* 3354 4062: contig of 710 bp in length
* 4063 4162: gap of 100 bp
* 4163 4901: contig of 739 bp in length
* 4902 5001: gap of 100 bp
* 5002 5741: contig of 740 bp in length
* 5742 5841: gap of 100 bp
* 5842 6615: contig of 774 bp in length
* 6616 6715: gap of 100 bp
* 6716 7472: contig of 757 bp in length
* 7473 7572: gap of 100 bp
* 7573 8264: contig of 692 bp in length
* 8265 8364: gap of 100 bp
* 8365 9124: contig of 760 bp in length
* 9125 9224: gap of 100 bp
* 9225 9981: contig of 757 bp in length
* 9982 10081: gap of 100 bp
* 10082 10774: contig of 693 bp in length
* 10775 10874: gap of 100 bp
* 10875 11611: contig of 737 bp in length
* 11612 11711: gap of 100 bp
* 11712 12449: contig of 738 bp in length

```

```
* 12550 12549: gap of 100 bp
* 12550 1302: contig of 753 bp in length
* 13303 13402: gap of 100 bp
* 13403 14163: contig of 761 bp in length
* 14164 14263: gap of 100 bp
* 14264 15035: contig of 772 bp in length
* 15036 15135: gap of 100 bp
* 15136 15905: contig of 770 bp in length
* 15906 16005: gap of 100 bp
* 16006 16769: contig of 764 bp in length
* 16770 17592: contig of 723 bp in length
* 17593 17692: gap of 100 bp
* 17693 18415: contig of 723 bp in length
* 18416 18515: gap of 100 bp
* 18516 19263: contig of 748 bp in length
* 19264 20132: contig of 769 bp in length
* 20133 20232: gap of 100 bp
* 20233 20965: contig of 733 bp in length
* 20966 21065: gap of 100 bp
* 21066 21792: contig of 727 bp in length
* 21793 21892: gap of 100 bp
* 21893 22603: contig of 711 bp in length
* 22604 22703: gap of 100 bp
* 22704 23447: contig of 744 bp in length
* 23448 23547: gap of 100 bp
* 23549 24290: contig of 743 bp in length
* 24291 24390: gap of 100 bp
* 24391 25162: contig of 772 bp in length
* 25163 25262: gap of 100 bp
* 25263 26044: contig of 782 bp in length
* 26045 26144: gap of 100 bp
* 26145 26921: contig of 777 bp in length
* 26922 27021: gap of 100 bp
* 27022 27737: contig of 716 bp in length
* 27738 27837: gap of 100 bp
* 27838 28597: contig of 760 bp in length
* 28598 28697: gap of 100 bp
* 28699 29448: contig of 751 bp in length
* 29449 29548: gap of 100 bp
* 29549 30318: contig of 770 bp in length
* 30319 30418: gap of 100 bp
* 30419 31169: contig of 751 bp in length
* 31170 31269: gap of 100 bp
* 31270 32005: contig of 736 bp in length
* 32006 32105: gap of 100 bp
* 32106 32854: contig of 749 bp in length
* 32855 32954: gap of 100 bp
* 32955 33714: contig of 760 bp in length
* 33715 33814: gap of 100 bp
* 33815 34591: contig of 777 bp in length
* 34592 34691: gap of 100 bp
* 34692 35448: contig of 757 bp in length
* 35449 35548: gap of 100 bp
* 35549 36316: contig of 768 bp in length
* 36317 36416: gap of 100 bp
* 36417 37139: contig of 723 bp in length
* 37140 37239: gap of 100 bp
* 37240 37996: contig of 757 bp in length
* 37997 38096: gap of 100 bp
* 38097 38839: contig of 743 bp in length
* 38840 38939: gap of 100 bp
* 38940 39714: contig of 775 bp in length
* 39715 39814: gap of 100 bp
* 39815 40483: contig of 669 bp in length
* 40484 40583: gap of 100 bp
* 40584 41307: contig of 724 bp in length
* 41308 41407: gap of 100 bp
* 41408 42110: contig of 703 bp in length
* 42111 42947: gap of 100 bp
* 42948 43047: gap of 100 bp in length
```

```
* 43048 43812 43911: contig of 764 bp in length
* 43812 43912 43911: gap of 100 bp
* 43912 44673 44673: contig of 762 bp in length
* 44674 44773 44773: gap of 100 bp
* 44774 45539 45539: contig of 766 bp in length
* 45540 45639 45639: gap of 100 bp
* 45640 46410 46410: contig of 771 bp in length
* 46411 46510 46510: gap of 100 bp
* 46511 47236 47236: contig of 726 bp in length
* 47237 47336 47336: gap of 100 bp
* 47337 48096 48096: contig of 760 bp in length
* 48097 48196 48196: gap of 100 bp
* 48197 48957 48957: contig of 761 bp in length
* 48958 49057 49057: gap of 100 bp
* 49059 49802 49802: contig of 745 bp in length
* 49803 49902 49902: gap of 100 bp
* 49903 50632 50632: contig of 730 bp in length
* 50633 50732 50732: gap of 100 bp
* 50733 51463 51463: contig of 731 bp in length
* 51464 51563 51563: gap of 100 bp
* 51564 52288 52288: contig of 725 bp in length
* 52289 52388 52388: gap of 100 bp
* 52389 53157 53157: contig of 769 bp in length
* 53158 53257 53257: gap of 100 bp
* 53258 54035 54035: contig of 778 bp in length
* 54036 54135 54135: gap of 100 bp
* 54136 54894 54894: contig of 759 bp in length
* 54895 54994 54994: gap of 100 bp
* 54995 55750 55750: contig of 756 bp in length
* 55751 55850 55850: gap of 100 bp
* 55851 56622 56622: contig of 772 bp in length
* 56623 56722 56722: gap of 100 bp
* 56723 57456 57456: contig of 734 bp in length
* 57457 57556 57556: gap of 100 bp
* 57557 58258 58258: contig of 702 bp in length
```

Query Match 84.0%; Score 16.8; DB 2; Length 71049;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACGATACATCG 20
DB 23922 AGTCGCAACGATACATCG 23941

RESULT 18
AC112029 2/c
WPCOMMENT
Sequence Split into 6 fragments LOCUS AC112029 Accession AC112029

Fragment Name Begin End
AC112029_0 1 110000
AC112029_1 100001 210000
AC112029_2 200001 310000
AC112029_3 300001 410000
AC112029_4 400001 510000
AC112029_5 500001 569567
Continuation (3 of 6) of AC112029 from base 200001 (AC112029 Rattus norvegicus clone CH2;

Query Match 84.0%; Score 16.8; DB 2; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACGATACATCG 20
DB 102858 AGTCGCAACGATACATCG 102839

RESULT 19
AC112029 3/c
WPCOMMENT
Sequence Split into 6 fragments LOCUS AC112029 Accession AC112029
Fragment Name Begin End
AC112029_0 1 110000

AC112029_1 100001 210000
AC112029_2 200001 310000
AC112029_3 300001 410000
AC112029_4 400001 510000
AC112029_5 500001 569567
Continuation (4 of 6) of AC112029 from base 300001 (AC112029 Rattus norvegicus clone CH2)

Query Match 84.0%; Score 16.8; DB 2; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGTGGCAACGTCATCG 20
Db 2858 AGCAGCAACGTCATCG 2839

RESULT 20
OSJN00288/c 112404 bp DNA linear PIN 10-FEB-2004
LOCUS OSJN00288 112404 bp DNA linear PIN 10-FEB-2004
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OJ000315_02,
complete sequence.

ACCESSION AL731639
VERSION AL731639.3 GI:38569166
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriatoidae; Oryzaceae; Oryza.

REFERENCE 1
AUTHORS Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,
Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Yang, K., Yu, S., Tang, Y.,
Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L., S., Yu, Z., Fan, D.,
Liu, X., Lu, Y., Li, C., Mu, Y., Sun, T., Lei, H., Li, F., Hu, H., Guan, J.,
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R.,
Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J.,
Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,
Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S.,
Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,
Li, J., Hong, G., Xue, Y., and Han, B.

TITLE Sequence and analysis of rice chromosome 4
JOURNAL Nature 420 (6913), 316-320 (2002)
MEDLINE 22337377
PUBMED 12447439
REFERENCE 2
AUTHORS Fu, G., Wang, S., Ren, S., X., Lv, G., Lin, W., Gu, W., Q., Zhu, G., F.,
Tu, Y., F., Jia, P., Yin, H., F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X., Y., Shao, Y., Sun, Y., Hu, Q., P., Zhang, X., L., Zhang, W.,
Wang, L., U., Ding, C., M., Sheng, H., H., Gu, J., L., Chen, S., T., Ni, L.,
Zhu, F., H., Han, B., Feng, Q., Huang, Y., C., Li, Y., Zhu, J., J., Zhao, Q.,
Hu, X., Liu, Y., L., Mu, Y., Yu, Z., Chen, L., Fan, D., L., Weng, Q., J.,
Zhang, L., Lu, Y., Q., Yu, S., L., Liu, X., H., Lu, T., T., Zhang, Y., J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P., X., Qian, Y., M., Yang, K.,
Zhou, B., Chen, Z., H., Hao, P., Zhang, L., Wu, M., Zhang, R., Q., Guan, J., P.
and Hong, G., F.

TITLE Direct Submission
JOURNAL Submitted (04-MAY-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OJ000315_02.
COMMENT On Dec 1, 2003 this sequence version replaced gi:32487896.
Web site: http://www.ncgr.ac.cn
Summary Statistics
Assembly program: phrap

REMARK This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Egenes (http://www.sotberry.com/),
genscan (http://CCR-081.mt.edu/GENSCAN.html), GenemarkHMM
(http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean
Bddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the

complete sequence against NCBI none redundant protein database (nr)
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES
source 1. 112404
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="4"
/clone="OJ000315_02"
/clone_lib="NonIntro-OUT"
complement (27..1985)
/gene="OJ000315_02.1"
complement (join(27..154,455..719,816..1985))
/gene="OJ000315_02.1"
/codon_start=1
/protein_id="CAE05356.1"
/db_xref="GI:32487897"
/db_xref="FBI:07618"
/translation="MVDEELFDKSSNDHSISSEEDMLVRSYNLNVSGYHGNISYOC
PSLDIDEDYDISPNKRLTNTMTSONSFCLSGAALISANFTLANNTCKGLIGETIL
PELDSFNSFRIVSSPMSRLDLISTQSGPVSIESIPEISKNIWASAPTVSSNF
LSTFEIKMAGAGEDRVOAQCSEKRMILCGIYDNGRDLADFLATLYDNIYVL
YLSECRIKOENGLGSPGSLNGVSKSLTAMEFAEMEDYKFSPTAGYLKCLTAV
EOANDPILCMWEOPMDDRPDLVSGCVTVVILHGTDLITLNGDSRAVAVASPSGM
DKRAVYLTEHSLLENPLEYQKLADHPNRPSPYMGKIKGKATRAVGVYLRQK
LNDALWILKRNLCSPYVYNTPHVSHRVTDDLFVVGSGGLDFPESNDVVOVL
YQFMHNPIDPAPYLLIEOLLKAKKAALTAELMKIPVGRSKYHDDVTITVILIG
NAQRTMTASTSL"
complement (12897..15028)
/gene="OJ000315_02.2"
complement (join(12897..13094,13461..13545,13645..13730,
14707..14731,14955..15028))
/gene="OJ000315_02.2"
/codon_start=1
/protein_id="CAE05357.3"
/db_xref="GI:38569167"
/translation="MTTDEFTGSGYSGICSPVACHVYNTYRIRHISTOWSGYDCLVLD
QRCDSFSTYIFKNVDSRPPRNLRDLRGLTPRANKALSDPEVTKDHENSIRHGRHLV
SMVHQRLLRRVQPMPLRLRLRLAFISAQEWLLDLDKEYGMKMKCTRRDKTS"
17220..17643
/gene="OJ000315_02.3"
join(17220..17290,17541..17643)
/gene="OJ000315_02.3"
/codon_start=1
/protein_id="CAE05358.3"
/db_xref="GI:38569168"
/translation="MASSSSSMQALSIMSCTIVYKTFSPCEPVKPKMKVYKVRBRK
LPVSPKFTARVP"
complement (24736..27684)
/gene="OJ000315_02.4"
complement (join(24736..25083,26698..26751,27465..27561,
27556..27684))
/gene="OJ000315_02.4"
/codon_start=1
/protein_id="CAE05359.3"
/db_xref="GI:38569169"
/translation="WAVEGEYSIMELKLPFRFDCKYKENSETLFMEDSWGSK
EGGRGMOGAQLRQYRKQGVMAENVYDAGEBRFGHGHSSGTTPTGAASPK
MRSWSAASASSGSGHSGPKVCAPATHAGSFRCRLRSSHGHSPSSPTAAAP
AAAVQSSSRTVAQ"
complement (28215..31978)
/gene="OJ000315_02.5"
complement (join(28215..28502,28743..28883,29291..30228,
30405..30644,30781..31978))
/gene="OJ000315_02.5"
/codon_start=1
/protein_id="CAE05360.3"
/db_xref="GI:38569170"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

/db_xref="TrEMBL:O7X7X7"
 /translation="MCCPRSGCVT.IOSAKGI.VCKEHNPLAITYT.IMMVGSIDLEFQALBE
 GPTQHFKNMWTMYKAMQEKGLFHRMPLINKKTLDELAFDSI.RKKGTS.VHYKVKKXMOE
 AEAVERLHEWTVVEGVDEKDYD.PHEHGSINIGVIEYDMWALKTDVYLRVGMN
 TLKSLPLMTLMTPLGLVYKAHFFLEQI.VELGMFRDCAQERVAVEVEQNEPENINM
 QKNKKPRKEEASNNRVLKQAGDKKKA.IVEEDSDNSAQAEIVKAKOMEDENLAR
 VOLNEHFVYKOIQLDKDVAGCMIEGGIREEAQAGEBNEKTKIOTEGMADKSGTSQV
 ENDSVEYIDDESPVQVGSSEENSOESDPAVAVGVLSQTNEDMTLEKKSILR
 LLEKREKCALPADDTNLT.IKNEELARINL.FLEKELNTSQKSCQVGVNLSIDIN
 IDEYHSDCMSDPDVAEMNT.I.SQSNSEKHPDICTOETVUD.FAKKDLIDGGG
 KPSNMWT.PKESGGMVYGTEDPMELEIVSKNPLFLVKKYKRNKPFVYRLITYYGA
 AQPERKEEFLLEBLQDCYDEP.LLIGDVFILQENKRNKGVNMS.IAFNSITEN
 ANLRMEALKEGKFTVANNKQSLTEKLDRLIFNKNWBI.KEPPLAGFVLEIYDHPPL
 LINVEGTEMLPELPEFENSMTLREDIN.VVQVWNSCNINGS.IDKXOMLQJMRKL
 KGMNMMNGMYKAKKHEIMEKIEDIPKACEYAGTILERKRGLEBELKKNLNEEL
 ITDEKMRKIVRELKLVGYITLIPKCKEARMQKRRPICLNAKGPEIOWDMILKVKV
 GKVA.IRNDQIQHYFTTHKGLRQGDPLSLFNIAADLTILIKRAEQGILKXGM
 DYMERGVAILDQADHTICLT"
 36050..36766
 /gene="OJ000315_02.6"
 join(36050..36423,36754..36766)
 /gene="OJ000315_02.6"
 /codon_start=1
 /protein_id="CAE05361.3"
 /db_xref="GI:38569171"
 /db_xref="TrEMBL:O7X7Y8"
 /translation="MERARRGSSGGGTGCGAAGHARGDDGSSAVATRCARNDG
 ATGMRRARRGGVAVAAVVGARRWTARRRNDGDDGAGLGAACGDDGVGGEAEAR
 RAGRENNAKGELGGLDVEHRRVYS"
 complement(40401..42688)
 /gene="OJ000315_02.7"
 complement(join(40401..40769,42056..42688))
 /gene="OJ000315_02.7"
 /codon_start=1
 /protein_id="CAE05362.3"
 /db_xref="GI:38569172"
 /db_xref="TrEMBL:O7X7G9"
 /translation="MSTSHLSSLSLILPSPSLILPGLPTRRVVGAGGELRRRR
 AMSAPLRLSTLEAKVSRBOELTASAGHGMGLSMURYCEVAPRKSGREYK
 ALPSARVRERLIPAREGTARAPSCASVSRERESVGVSAPSTLLARLSAYTF
 DDPSPQYTKAKRHTLSPHPRAGTDLARSMASASENSVTSAAAMRRRTJACGL
 ERLMSKMYVPSAKITKPSRREBEVDLVWATIRFGRRLSLAAMSRMARWAPRQ
 YRGRGEBEENCDDGMVGQVGTAYEVVRAKELRGSDRAADEVAEEBAVVC"
 complement(43876..45030)
 /gene="OJ000315_02.8"
 complement(join(43876..44247,44336..44561,44657..45030))
 /gene="OJ000315_02.8"
 /codon_start=1
 /protein_id="CAE05363.3"
 /db_xref="GI:38569173"
 /db_xref="TrEMBL:O7X653"
 /translation="VAAVPTAVVAAVALVCGGGGARADQMDNEMARNRPFGGTG
 GGVVGLIPQSDVLEPPLNLEYLFAFPQMSALIGLDISLIGGGAPAFQATQ
 AALPFPVDIAIQFCQGVGHLEIKOVNRPPLPDLISANGLKIVESADAVTPRD
 PNPENPENLFLASVYII.PYVGLTGYVANSRLLTPQAKI.VWAGLLEVAADAVIRA
 LLYHGSLRVSATYGVGAF.LRHSIELANVYGRGVNDEGIVAPGGGPEGQYGNII
 AGDRSLAYDRTPTEIILGVYVSGDPAPAKGFFQAGAGR.IARAFIA"
 complement(46275..46673)
 /gene="OJ000315_02.9"
 complement(46275..46673)
 /gene="OJ000315_02.9"
 /codon_start=1
 /protein_id="CAE05364.3"
 /db_xref="GI:38569174"
 /db_xref="TrEMBL:O7X7G5"
 /translation="MSEHAVVEDVVRKILLSAKGAGVAINAALLLEISFPAALKC
 KKEEESAPLARVPERKIGTSGLVAKKEERLYRRTRVALVTRPTEKITIPDTRKLN
 VEHRIPNKKKLGDDNDEGADVRNDPVVC"
 51378..51761
 /gene="OJ000315_02.10"
 51378..51761
 /gene="OJ000315_02.10"
 /gene="OJ000315_02.10"

Query Match	Similarity	Score	DB	Length
Best local	90.0%	Pred. No. 3.1e+02;		112404;
Matches	18;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;

QY	1	ACTCGCAACAGTACAAATGG	20
Db	111288	AGTACCAACAGTACAAATGG	111269

RESULT 21	AL119153/c	115595 bp	DNA	linear	HTG 10-UTL-2001
LOCUS	AL119153				
DEFINITION	Homio sapiens chromosome 1 clone RP4-753A11, 10 unordered pieces.				
ACCESSION	AL119153				
VERSION	AL119153.4 GI:9863525				
KEYWORDS	HTG; HTGS; PHASE1; HTGS_CANCELLED.				
SOURCE	Homio sapiens (human)				
ORGANISM	Homio sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homio.				

REFERENCE	Plumb, B.
AUTHORS	Plumb, B.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
	requests: clonerequests@sanger.ac.uk
	On Aug 21, 2000 this sequence version replaced gi:9212193.
	----- Genome Center
	Center: Sanger Centre
	Center code: SC
	Web site: http://www.sanger.ac.uk
	Contact: humquerry@sanger.ac.uk
	----- Project Information
	Center project name: dt753A11
	----- Summary Statistics
	Assembly program: XGAP; version 4.5
	Sequencing vector: plasmid; 108752; 100% of reads
	Chemistry: Dye-terminator Big Dye; 100% of reads
	Consensus quality: 110424 bases at least Q40
	Consensus quality: 112822 bases at least Q30
	Consensus quality: 114272 bases at least Q20
	Insert size: 115685; sum-of-contigs
	Insert size: 127770; 7.3% error; agarose-fp
	Quality coverage: 3.64x in Q20 bases; sum-of-contigs Quality
	coverage: 3.37x in Q20 bases; agarose-fp

	* NOTE: This is a 'working draft' sequence. It currently
	* consists of 10 contigs. The true order of the pieces
	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	1 13312: contig of 13212 bp in length
	* 13213 gap of 100 bp
	* 13313 20141: contig of 6829 bp in length
	* 20142 gap of 100 bp
	* 20242 20470: contig of 9229 bp in length
	* 20471 29570: gap of 100 bp
	* 29571 35519: contig of 5949 bp in length
	* 35520 35619: gap of 100 bp
	* 35620 52338: contig of 16719 bp in length
	* 52339 gap of 100 bp
	* 52339 88864: contig of 36526 bp in length
	* 88965 89064: gap of 100 bp
	* 89065 97388: contig of 8324 bp in length
	* 97389 97489: gap of 100 bp
	* 97489 108993: contig of 11505 bp in length
	* 108994 109093: gap of 100 bp
	* 109094 113951: contig of 4858 bp in length
	* 113952 114051: gap of 100 bp
	* 114052 116585: contig of 2534 bp in length.

```

FEATURES
Source      Location/Qualifiers
1. 116585
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP4-753A11"
    /clone_lib="RPCI-4"
    1. 13212
    /note="assembly_fragment:01103
    fragment_chain:1"
    1313. 20141
    /note="assembly_fragment:00150
    fragment_chain:1"
    20242. 29470
    /note="assembly_fragment:00511.0"
    29571. 35319
    /note="assembly_fragment:00511.1"
    35620. 52338
    /note="assembly_fragment:00628"
    52439. 88964
    /note="assembly_fragment:00071
    fragment_chain:2"
    89065. 97388
    /note="assembly_fragment:00168
    fragment_chain:2"
    97489. 108993
    /note="assembly_fragment:00346
    fragment_chain:2"
    109094. 119951
    /note="assembly_fragment:00939
    fragment_chain:2"
    114052. 116585
    /note="assembly_fragment:00581
    fragment_chain:2
    clone_end:SP6
    vector_side:right"

ORIGIN
Query Match      84.0%; Score 16.8; DB 2; Length 116585;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy      1 AGTCGCAACACGACACATGC 20
Db      52779 AATCAGCAACAGACATGC 52760

RESULT 22
AC006945      120311 bp      DNA      linear      ROD 17-SNP-2003
LOCUS      AC006945
DEFINITION      Mus musculus clone ct7-453113 map 6 strain 129/Sv, complete
sequence.
AC006945
VERSION      AC006945.17      GI:34787443
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Mus musculus Chromosome 6 BAC Clone ct7-453113 In CES Region
Unpublished
2 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (02-MAR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission

```

```

JOURNAL
Submitted (07-JUL-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (24-JUL-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (15-FEB-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (22-FEB-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
7 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (15-JUN-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
8 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (18-JAN-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
9 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (25-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
10 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
11 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Sep 17, 2003 this sequence version replaced gi:24137451.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
1..120311
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/map="6"
/clone="cct7-453113"
/clone_11b="ClbCC7 mouse BAC library"
ORIGIN
Query Match 84.0%; Score 16.8; DB 10; Length 120311;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
1 AGTCGCAACAGTACATGG 20

```


DEFINITION	OSUNBA0032E21, complete sequence.
ACCESSION	AF377947
VERSION	AF377947.3
KEYWORDS	GI:21070919
SOURCE	RTG.
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	Oryza sativa (japonica cultivar-group)
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE	1 (bases 1 to 137580)
REFERENCE	Bestman,A.P., Smith,S.C., Bertin,N., Liang,C., Najjar,F.Z., Pratt,L.H. and Cordonnier-Pratt,M.-M.
REFERENCE	2 (bases 1 to 137580)
REFERENCE	Bestman,A.P., Smith,S.C., Gingle,A., Pratt,L.H. and Cordonnier-Pratt,M.-M.
REFERENCE	Direct Submission
REFERENCE	Submitted (01-MAY-2001) Botany, University of Georgia, Miller Plant Sciences, Athens, GA 30602, USA
REFERENCE	3 (bases 1 to 137580)
REFERENCE	Bestman,A.P., Smith,S.C., Bertin,N., Liang,C., Najjar,F.Z., Pratt,L.H. and Cordonnier-Pratt,M.-M.
REFERENCE	Direct Submission
REFERENCE	Submitted (22-MAY-2002) Botany, University of Georgia, 3604 Miller Plant Sciences, Athens, GA 30606, USA
REFERENCE	Sequence update by submitter
REFERENCE	On May 22, 2002 this sequence version replaced gi:15384618.
REFERENCE	The orientation of the sequence is from M13-21 to Jenvrev of the BAC clone. The assembly has been confirmed via restriction digest.
REFERENCE	Genes were predicted from the integrated results of the following: GENSCAN1.0, GenMark 2.2. The genomic sequence and predicted gene database, SWISSPROT, and NCBI EST database. Protein similarities of the coding regions were searched against SWISSPROT with BLASTp.0. Genes are annotated based on the level of evidence. Genes with high homology hits to known proteins were annotated as the name of the protein hit. Genes with no protein data but high EST homology were annotated as unknown proteins. Predicted proteins are genes predicted over most of their length by both GenScan and GeneMark.
REFERENCE	location/Qualifiers
REFERENCE	1. .137580
REFERENCE	/organism="Oryza sativa (japonica cultivar-group)"
REFERENCE	/mol_type="genomic DNA"
REFERENCE	/db_xref="taxon:39947"
REFERENCE	/chromosome="3"
REFERENCE	/clone="OSUNBA0032E21"
REFERENCE	/clone_1b="BAC"
REFERENCE	1266..11420
REFERENCE	/rpt_unit="cgg"
REFERENCE	1749..1925
REFERENCE	/rpt_unit="cgg"
REFERENCE	complement(2314..2374)
REFERENCE	/rpt_unit="cgg"
REFERENCE	complement(2846..2902)
REFERENCE	/rpt_unit="cgg"
REFERENCE	5573..7073
REFERENCE	/note="High similarity BLASTN hit to LTR of Oryza sativa PAC clone P0416G11"
REFERENCE	complement(9044..9230)
REFERENCE	/rpt_unit="cgg"
REFERENCE	complement(14921..15083)
REFERENCE	/rpt_unit="ccccg"
REFERENCE	<16521..>18404
REFERENCE	/gene="OSUNBA0032E21.01"
REFERENCE	join(<16521..17973,18307..>18404)
REFERENCE	/gene="OSUNBA0032E21.01"
REFERENCE	/product="unknown protein"
REFERENCE	join(16521..17973,18307..18404)
REFERENCE	/gene="OSUNBA0032E21.01"
REFERENCE	/note="confirmed by high quality EST hit to GB acc_num AU0031009"

Query Match	Best Local Similarity	90.0%;	Score 16.8;	DB 8;	Length 137580;
Matches	18; Conservative	0;	Mismatches	2;	Indels
					Gaps
QY	1 AGTCGCAACAGTACATGG 20				
Db	18241 AGTAGCAACCAATACATGG 18222				
RESULT 25					
LOCUS	AC104996	155848 bp	DNA	linear	PII 31-MAR-2002
DEFINITION	Homo sapiens chromosome 17, clone RP11-114804, complete sequence.				
ACCESSION	AC104996				
VERSION	AC104996.4	GI:19849376			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				


```

repeat_region      /rpt_family="L2"
                    11971..12186
                    /rpt_family="Tigerza"
repeat_region      12187..12266
                    /rpt_family="TTCOR"
repeat_region      complement(12311..12594)
                    /rpt_family="Aludo"
repeat_region      12596..12632
                    /rpt_family="Tigerza"
repeat_region      12678..12914
                    /rpt_family="AT_rich"
repeat_region      complement(13242..13576)
                    /rpt_family="Alut"
repeat_region      13780..13881
                    /rpt_family="LIME3A"
repeat_region      complement(13888..14298)
                    /rpt_family="MLT1A1"
repeat_region      14303..14610
                    /rpt_family="LIME3A"
repeat_region      14953..15244
                    /rpt_family="L1P"
repeat_region      15245..15608
                    /rpt_family="THE1B"
repeat_region      15609..17164
                    /rpt_family="THE1B-int"
repeat_region      17165..17528
                    /rpt_family="THE1B"
repeat_region      17529..18050
                    /rpt_family="L1P"
repeat_region      18028..18163
                    /rpt_family="LIME3"
repeat_region      complement(18164..18586)
                    /rpt_family="MERS7A"
repeat_region      18587..18721
                    /rpt_family="LIME3"
repeat_region      18750..18780
                    /rpt_family="(GA)n"
repeat_region      complement(18811..19233)
                    /rpt_family="L1"
repeat_region      19234..19947
                    /rpt_family="L1P10"
repeat_region      complement(19948..20072)
                    /rpt_family="L1"
repeat_region      complement(20088..22341)
                    /rpt_family="L1"
repeat_region      complement(22361..22521)

```

Query Match 84.0%; Score 16.8; DB 9; Length 155848;
 Best Local Similarity 90.0%; Pred. No. 3.1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
 Db 13683 AGTCACAAACAGTACAGG 13702

RESULT 26
 HHV6AGNM 159321 bp DNA linear VRL 17-FPB-1997
 LOCUS Human herpesvirus-6 (HHV-6) U1102, variant A DNA, complete virion
 DEFINITION genome.
 ACCESSION X83413.1
 VERSION X83413
 KEYWORDS alkali exonuclease; complete genome; DNA polymerase; DR2; DR3; DR4; DR5; DR6; DR7; DR8; dUMPase; ganiclovir kinase; GCR; GH; GU; glycoprotein B; glycoprotein gp82/105; glycoprotein H; glycoprotein L; glycoprotein W; gm; helicase/primase complex; IE glycoprotein; I1; integral membrane protein; large tegument protein; L1; L11; major capsid protein; major DNA binding protein; membrane protein; origin binding protein; pDR1; phosphotransferase; pol processivity factor; pP100; pP41; replication protein; R1; transactivator; U1; U10; U100; U11; U12; U13; U14; U15; U16; U17; U18; U19; U20; U21; U22; U23; glycoprotein; U24; U25; U26; U27; U28; U29; U3; U30;

SOURCE
 ORGANISM
 Human herpesvirus 6
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 Betaherpesvirinae; Roseolovirus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 89279291
 2543772
 Lawrence,G.L., Chee,M., Craxton,M.A., Gompels,U.A., Honess,R.W. and
 Barrett,B.G.
 Human herpesvirus 6 is closely related to human cytomegalovirus
 J. Virol. 64 (1), 287-299 (1990)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 90060132
 2152817
 Little,E., Lawrence,G., Liu,M.Y., Barrett,B.G. and Arrand,J.R.
 Identification, cloning, and expression of the major capsid protein
 gene of human herpesvirus 6
 J. Virol. 64 (2), 714-722 (1990)

JOURNAL
 MEDLINE
 PUBMED
 90112641
 2153237
 Martin,M.E., Thomson,B.J., Honess,R.W., Craxton,M.A., Gompels,U.A.,
 Liu,M.Y., Little,E., Arrand,J.R., Teo,I. and Jones,M.D.
 The genome of human herpesvirus 6: maps of unit-length and
 concatemeric genomes for nine restriction endonucleases
 J. Gen. Virol. 72 (Pt 1), 157-168 (1991)

JOURNAL
 MEDLINE
 PUBMED
 9116306
 1846644
 Thomson,B.J., Efsthion,S. and Honess,R.W.
 Acquisition of the human adeno-associated virus type-2 rep gene by
 human herpesvirus type-6
 Nature 351 (6321), 76-80 (1991)

JOURNAL
 MEDLINE
 PUBMED
 9126542
 1851252
 Chang,C.K. and Balachandran,N.
 Identification, characterization, and sequence analysis of a cDNA
 encoding a phosphoprotein of human herpesvirus 6
 J. Virol. 65 (6), 2884-2894 (1991)

JOURNAL
 MEDLINE
 PUBMED
 91237802
 1851860
 Teo,I.A., Griffin,B.E. and Jones,M.D.
 Characterization of the DNA polymerase gene of human herpesvirus 6
 J. Virol. 65 (9), 4670-4680 (1991)

JOURNAL
 MEDLINE
 PUBMED
 91333007
 1651403
 Martin,M.E., Nicholas,J., Thomson,B.J., Newman,C. and Honess,R.W.
 Identification of a transactivating function mapping to the
 putative immediate-early locus of human herpesvirus 6
 J. Virol. 65 (10), 5381-5390 (1991)

JOURNAL
 MEDLINE
 PUBMED
 91374590
 1654446
 Gompels,U.A., Carrs,A.L., Sun,N. and Arrand,J.R.
 Infectivity determinants encoded in a conserved gene block of human
 herpesvirus-6
 DNA Seq. 3 (1), 25-39 (1992)

JOURNAL
 MEDLINE
 93091256


```
complement(join(25712, .25917, 26704, .27220))
CDS
```

```

/gene="OBGNA0042115.5"
/codon_start=1
/protein_id="CAE04883.2"
/db_xref="GI:38347592"
/db_xref="TrEMBL:P7XLJY5"
/translation="MPPLTPIHIEKGVGRSQROQBSASAAAAAAGGOLLQCRP
GSGASAFHGRGTGEPHMLQARORPTQOCDLILAGVKGATATASVPAAGGCELEPFA
AGRTSSKYLIVAVORLIMPLHVMASAAASVADLVAAVAVLYIEGGRPPLPADSS
DFELHYSQSLSDLEPRKXWELDEKRSFPLCPRSSAAVHAPESSDEASRIRRDAA
PAARAGCAPAWVSVMGFWPMW"
complement(47475..5286)
/gene="OSU2NA042115.6"
complement(1011(47475)..49504,49559..50502,50555..50583,

```

```

50658.1..50861.50874.50950.51188.51589.51717.52285.
52349.52861.
/ gene="OSUNa0042115.6"
/ codon_start=1
/ protein_id="CAE04884.2"
/ db_xref="GI:38347593"
/ db_xref="nt:EMBL:07X1.Y4"
/ translation="VKHLISLHSLINKVYCTGVDPFHDMELTSEQQLHRAQAS
NALISTSLSEFNKVDGLSEAKELCDTLQAHGSPAVAVBAKLELSERLDLRFWMDIKR
ETQEWYQDMMLIVKRIKGLSEDMTTHVKKVLLAEFGPNKPLDVSIVIRKKQFKEL
TSLDIQYR.VSHEMQEBEARSATRRVRVRFHPIRKSGYKQKPDVKKQSKQKACN
CGCYGFIADPKNSAKAKGKKCPERPAVAADNHWAEWYEGSDEEVEVQKPRPD
KVDGEGVATVTFKSSSSSKRLFNINSLDDDSHYSCFMAQSGKQVTOQKPSLSD
VSDSEESNEMLDYLKSPSPMOHLAKMLRALDITLKENRLGLTVKYGCMYTHYS
KDLAYTTIHAHNRKHGGLGKSPSSPVKRYVMNGECLMIPRESKVAIAPENNVAD
EYHTCWVWKALAYSTDGSSWVDSCGTTNMTGERSMFSLDEKSGSENVTFPGDDCK
EKQIITIRSVIRDDSSIIAFKGVALKGDIYLVDDVAVPAECLAFKSIKMMWHLARL
AHGMMNLSLILKGGHILGNSVFPENDRVCSAQGKQVSGFLHAKINIMTTTPEL

```

Query Match	84.0%;	Score 16.8;	DB 2;	Length 159683;
Best Local Similarity	90.0%;	Pred. No. 3.1e+02;		
Matches 18;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

```

QY      1 AGTCGCAACAGTACAATGG 20
          ||| ||||| ||||| |||
Db     14514 AGTAGCAAAACAGTACAATTG 14495

```

RESULT 28	
LOCUS	AL359894
DEFINITION	AL359894 161040 bp DNA linear PRI 21-MAY-2001
ACCSSION	Human DNA sequence from clone Rpl1-179c1 on chromosome 1. Contains
VERSION	AL359894
KEYWORDS	STSS, GSSS and a Cpg island, complete sequence.
SOURCE	AL359894.9 GI:10120085
ORGANISM	HTG, Cpg island.
	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 161040)
AUTHORS	Tromans, A.
TITLE	Direct Submision
JOURNAL	Submitted (21-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
	CB10 1SQ, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
	request: clonerequest@sanger.ac.uk
COMMENT	On Sep 13, 2000 this sequence version replaced gi:10086142.

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/RGP/Chr1>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP1-379C1 is from the library RPc1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-379C1. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-379C1 is at 1 in this sequence. The true left end of clone RPS-944F3.1 is at 10941 in this sequence. The true left end of clone RP4-753A11 is at 87987 in this sequence.

Location/Qualifiers

1..161040

```

source
Location/Qualifiers
1. .161040
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="SP1-1378C1"
/clone_1b="RPCT-11.2"
1140. .106
repeat_region
/note="MPR34 repeat: matches 3. .173 of consensus"
misc_feature
/note="match: GSS: Em:AQ201473"
6175. .6644
/note="match: GSS: Em:AQ201473"
misc_feature
/note="match: GSS: Em:AQ164294"
6182. .6629
/note="match: GSS: Em:B92723"
misc_feature
/note="match: GSS: Em:AQ757617"
9733. .9784
/note="13 copies 4 mer tata 96% conserved"
9788. .11493
/note="11P45 repeat: matches 4437. .6143 of consensus"
12597. .13032 repeat: matches 4437. .6143 of consensus"
18170. .18671
/note="18T2F repeat: matches 1. .450 of consensus"
19378. .19471
/note="47 copies 2 mer tt 66% conserved"
20396. .21033
/note="19 copies 2 mer aa 78% conserved"
/note="complement(21188. .23787) matches 1. .450 of consensus"
23670. .24131
/note="match: GSS: Em:AQ81193"
misc_feature
/note="match: GSS: Em:AQ692063"
23776. .23959
/note="match: GSS: Em:AQ90507"
28255. .28296
/note="21 copies 2 mer tt 75% conserved"
/note="complement(31696. .32086) matches 1. .450 of consensus"
33052. .33147
/note="match: GSS: Em:AQ806367"
repeat_region
/note="11P47 repeat: matches 6048. .6143 of consensus"
complement(33433. .34063)
/note="match: GSS: Em:AQ40019"
33711. .34151
/note="match: GSS: Em:AQ496221"
repeat_region
33372. .34384
/note="11MD repeat: matches -15. .405 of consensus"
36841. .36656
/note="28 copies 2 mer tt 71% conserved"
misc_feature
complement(37051. .37422)
/note="match: GSS: Em:AQ160500"
41667. .42036

```

```
repeat_region /note="LTR2B repeat: matches 3. .404 of consensus"
42057. .42212
repeat_region /note="3 copies 52 mer 84% conserved"
42058. .42237
repeat_region /note="30 copies 6 mer tatata 78% conserved"
42114. .42233
repeat_region /note="30 copies 4 mer tata 87% conserved"
42236. .42291
misc_feature /note="LTR2B repeat: matches 394. .448 of consensus"
42053. .42470
/note="match: GSS: Em:AQ827367"
misc_feature /note="match: GSS: Em:AQ150705"
44413. .44860
repeat_region /note="Tigerg2a repeat: matches 2. .429 of consensus"
43405. .45438
repeat_region /note="17 copies 2 mer tt 82% conserved"
47212. .47647
repeat_region /note="MER21B repeat: matches 339. .789 of consensus"
47695. .48392
repeat_region /note="MER50 repeat: matches 1. .722 of consensus"
48393. .48643
repeat_region /note="MER21B repeat: matches 78. .329 of consensus"
50121. .50192
repeat_region /note="36 copies 2 mer at 91% conserved"
50941. .51315
repeat_region /note="THERIC repeat: matches 1. .371 of consensus"
54119. .54164
/note="MER83-internal repeat: matches 3088. .3134 of
consensus"
54264. .54339
repeat_region /note="MER83-internal repeat: matches 2896. .3088 of
consensus"
54624. .54663
repeat_region /note="20 copies 2 mer tt 77% conserved"
55288. .55355
repeat_region /note="18 copies 4 mer cata 75% conserved"
55303. .55356
repeat_region /note="9 copies 6 mer acacac 79% conserved"
56737. .57186
misc_feature /note="match: GSS: Em:AQ834443"
56794. .56943
repeat_region /note="75 copies 2 mer aa 58% conserved"
56856. .57167
misc_feature /note="match: GSS: Em:AQ694480"
57634. .57815
repeat_region /note="91 copies 2 mer aa 56% conserved"
57843. .57903
repeat_region /note="LIM4A repeat: matches 6215. .6273 of consensus"
58125. .58409
misc_feature /note="match: GSS: Em:AQ392809"
complement(58148. .58503)
misc_feature /note="match: STS: Em:G06380"
58224. .58502
/note="match: STS: Em:G11051 Em:G11052"
58404. .58429
repeat_region /note="13 copies 2 mer aa 100% conserved"
60042. .60259
repeat_region /note="LIP repeat: matches 2859. .3076 of consensus"
60261. .60765
/note="LIPAS repeat: matches 5641. .6145 of consensus"
61895. .61942
repeat_region /note="24 copies 2 mer tt 81% conserved"
61896. .61943
repeat_region /note="12 copies 4 mer ttgt 83% conserved"
68755. .69139
repeat_region /note="MSTC repeat: matches 1. .405 of consensus"
70484. .70531
repeat_region /note="24 copies 2 mer tt 81% conserved"
74389. .75634
misc_feature /note="LTR9 repeat: matches 10. .625 of consensus"
complement(75294. .75722)
/note="match: GSS: Em:AQ039479"
```

```
repeat_region 8182. .83257
/note="LIP repeat: matches 1656. .3746 of consensus"
repeat_region 83243. .84833
/note="LIPa10 repeat: matches 4086. .5694 of consensus"
repeat_region 85449. .85939
/note="LIPa10 repeat: matches 5698. .6165 of consensus"
misc_feature complement(88001. .88737)
/note="match: GSS: Em:AQ253332"
complement(88277. .88711)
/note="match: GSS: Em:AQ622463"
90455. .90542
repeat_region /note="22 copies 4 mer gtgt 77% conserved"
90485. .90544
repeat_region /note="30 copies 2 mer gt 86% conserved"
90767. .92294
/note="Cpg island"
misc_feature /evidence=not_experimental
91748. .91795
repeat_region /note="8 copies 6 mer cgcgc 81% conserved"
91751. .91804
repeat_region /note="18 copies 3 mer cgc 79% conserved"
92549. .92576
repeat_region /note="14 copies 2 mer ac 96% conserved"
93377. .93410
/note="17 copies 2 mer aa 85% conserved"
```

Query Match 84.0% Score 16.8; DB 9; Length 161040;
Best Local Similarity 90.0%; Pred. No. 3,1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 AGTCGCAACAGTACATCG 20
Db 152329 AATCACAACAGTACATCG 152348

RESULT 29
AB021506 161573 bp DNA linear VRL 24-OCT-2000
LOCUS Human herpesvirus 6 DNA, variant B, complete genome, strain:HST.
DEFINITION AB021506
ACCESSION AB021506
VERSION AB021506.1 GI:4995977
KEYWORDS
SOURCE Human herpesvirus 6
ORGANISM Human herpesvirus 6
Virus; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Roseolovirus.
REFERENCE
1 (sites)
Isegawa,Y., Mukai,T., Nakano,K., Kagawa,M., Chen,J., Mori,Y.,
and Yamashita,K.
Sunagawa,T., Kawanishi,K., Sashihara,J., Hata,A., Zou,P., Kosuge,H.
Comparison of the complete DNA sequences of human herpesvirus 6
variants A and B
J. Virol. 73 (10), 8053-8063 (1999)
JOURNAL MEDLINE 99412319
PUBMED 10482554
REFERENCE 2 (bases 1 to 161573)
AUTHORS Isegawa,Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1998) Yuji Isegawa, Osaka University Medical
School, Department of Microbiology; Ymada-Oka 2-2, Suita, Osaka
565-0871, Japan (E-mail:iisegawa@micro.med.osaka-u.ac.jp,
Tel:81-6-879-3323, Fax:81-6-879-3329)
FEATURES
source location/Qualifiers
1..161573
/organism="Human herpesvirus 6"
/virus
/mol_type="genomic DNA"
/strain="HST"
/db_xref="taxon:10368"
/pop_variant="B"
complement(18. .365)
gene /gene="UL1"
complement(18. .365)
CDS
```

[illegible]

CDS complement (9467, 10768)
/gene="U2"
/note="92% identical to U2 of strain U102 of HHV-6
US22 gene family"
/codon_start=1
/protein_id="BA078232.1"
/db_xref="GI:4995989"
/translation="MCSQPCGASVHFHFRGRREYDSTLFVNSCSVLNIVDYVR
LPTVPLSFEFFFIILILSRIPRKKMPCSPPLGISMSLAAALCPSSCSFASGRD
LRCDAAVEYKMTAFVATLVARPLSADVDKDFSTAFHCRALSMVGAELRFASDM
IGITQAEALATSRGCGCGMGLTIGVITFAVSALIVGECGEIYAFNGVDALY
RLBDAFGMLKHLREREPYVSGKLMETGASFGMGVDALAAVAFVADALYLP
WBGAPPEFAVPRRARRRRLIRGCGAVAVIGSPRGVTLPLLRORVLMONQVY
AASLDGAVIRLDSFRALFAMGVRKLFXKRRPPEHLTMQLPTTCVAFVINDPAV
YQLSPMVEREKSAVSCGASTVYRDCEDTLRDGAGVDTS"
/complement (10891, 12051)
/gene="U3"
complement (10891, 12051)
CDS complement (10891, 12051)
Query Match 84.0%; Score 16.8; DB 14; Length 161573;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATGG 20
Db 74271 AGTCGCAACACATGATGG 74290
RESULT 30
AF157706 162114 bp DNA linear VRL 15-OCT-2003
LOCUS Human herpesvirus 6B strain Z29, complete genome.
ACCESSION AF157706.1 L13162 L14772 L16947
AF157706.1 GI:5733510
VERSION
KEYWORDS
SOURCE Human herpesvirus 6B
ORGANISM Human herpesvirus 6B
Human herpesvirus 6B
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Roseoloviruses.
REFERENCE
AUTHORS 1 (bases 19696 to 22886)
Pellett, P.E., Sanchez-Martinez, D., Dominguez, G., Black, J.B.,
Anton, E., Greenmoyer, C., and Dambaugh, T.R.
TITLE A strongly immunoreactive vitron protein of human herpesvirus 6
variant B strain Z29: identification and characterization of the
gene and mapping of a variant-specific monoclonal antibody reactive
epitope
JOURNAL Virology 195 (2), 521-531 (1993)
MEDLINE 93331710
PUBMED 7687803
REFERENCE 2 (bases 64106 to 84963)
Stamey, F.R., Dominguez, G., Black, J.B., Dambaugh, T.R., and
Pellett, P.E.
TITLE Intragenomic linear amplification of human herpesvirus 6B orlYLT
suggests acquisition of orlYLT by transposition
JOURNAL J. Virol. 69 (1), 589-596 (1995)
MEDLINE 95074921
PUBMED 7983761
REFERENCE 3 (bases 64106 to 84963; 106012 to 126166)
Lindquester, G.J., Inoue, N., Allen, R.D., Castell, J.W., Stamey, F.R.,
Dambaugh, T.R., O'Brian, J.J., Danovich, R.M., Frenkel, N., and
Pellett, P.E.
TITLE Restriction endonuclease mapping and molecular cloning of the human
herpesvirus 6 variant B strain Z29 genome
JOURNAL Arch. Virol. 141 (2), 367-379 (1996)
MEDLINE 96195263
PUBMED 8634027
REFERENCE 4 (bases 64106 to 84963)
Lindquester, G.J., O'Brian, J.J., Anton, E.D., Greenmoyer, C.A.,
Pellett, P.E., and Dambaugh, T.R.
TITLE Genetic content of a 20.9 kb segment of human herpesvirus 6B strain
Z29 spanning the homologs of human herpesvirus 6A genes U40-57 and
containing the origin of DNA replication
JOURNAL Arch. Virol. 142 (1), 103-123 (1997)

MEDLINE 97300856
PUBMED 9155876
REFERENCE 5 (bases 106012 to 126166)
Lindquester, G.J., Greenmoyer, C.A., Anton, E.D., O'Brian, J.J.,
Pellett, P.E., and Dambaugh, T.R.
TITLE Comparison of a 20 kb region of human herpesvirus 6B with other
human beta herpesviruses reveals conserved replication genes and
adjacent divergent open reading frames
JOURNAL Arch. Virol. 142 (1), 193-204 (1997)
MEDLINE 97300864
PUBMED 9155884
REFERENCE 6 (bases 1 to 162114)
Dominguez, G., Dambaugh, T.R., Stamey, F.R., Dewhurst, S., Inoue, N., and
Pellett, P.E.
TITLE Human herpesvirus 6B genome sequence: coding content and comparison
with human herpesvirus 6A
JOURNAL J. Virol. 73 (10), 8040-8052 (1999)
MEDLINE 99412318
PUBMED 10482553
REFERENCE 7 (bases 19696 to 22886)
Pellett, P.E.
TITLE Direct Submission
AUTHORS Submitted (26-JUL-1993) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REFERENCE 8 (bases 64106 to 84963; 106012 to 126166)
Pellett, P.E.
TITLE Direct Submission
AUTHORS Submitted (02-OCT-1993) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REFERENCE 9 (bases 1 to 162114)
Pellett, P.E., Dominguez, G., Dambaugh, T.R., Stamey, F.R., Dewhurst, S.
and Inoue, N.
TITLE Direct Submission
AUTHORS Submitted (10-JUN-1999) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REMARK Sequence updated by submitter
COMMENT On or before Aug 16 1999 this sequence version replaced gi:405170,
gi:405156, gi:303397.
FEATURES
source Location/Qualifiers
1..162114
/organism="Human herpesvirus 6B"
/mol_type="genomic DNA"
/strain="Z29"
/db_xref="taxon:32604"
/note="variant B"
1..8793
/rpt_type=direct
59..353
/note="heterogeneous telomeric unit"
/rpt_type=direct
/rpt_unit="tarsyc"
583..3988
/gene="U2"
/note="US22 gene family"
join(583..841,955..2975)
/gene="U3"
/codon_start=1
/product="pRI"
/protein_id="AA049614.1"
/db_xref="GI:5733511"
/translation="MPLTARAGHTLRLPLSHYWWLLIGRSLRHVSHYLRKGLRL
PLPWPBQCLHHPKPYKFLRYPCLTROPHLQGWPADSILMDFPXPYHPSADSKLL
PLGLITSAFMSRVSPTHCSEFHAADPSLMTGSSPMVLADLPAGSLFCHDVPFG
RLYLISVSFLFTGTGROCAIYRVLNRYRPLPSLWTCGDPPTATPLKLARTAR
GLAAPPILMRHLGSRSLSPYLSEWERTLVYLDPRRGMCCTLLSGSBSQSSSD
AGDTVMNATEKAAAGKBAERGGQCATDRIASPHLTRGRDCCRSLSQGEPSADEF
ACRCPDLDETCGEGGMLPFAVLNAPVLCILITGGLVLAIDLNDTSLMRADLELL
LRSLSLILSLGRLPLRPSSGSGAARPGYKEGSGRATATATATATATATATATATAT
GVTEKGRVTDVPSAHPSESEQTDHNGQESGHDGSDGSDGRGHDGARGRHAD
ETEPQGRHEDGQTDGSRDPAQSEVAVARDKXTGEGGSGRSGRAVQTVYGRGE


```

TITLE
JOURNAL
COMMENT
REFERENCE
AUTHORS
Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKenna, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, J., Murphy, T., Naylor, J., Nguyen, C.,
Nicoli, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thoman, N.,
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164107)
Barron, B., Linton, L., Nusbbaum, C., Landers, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, I.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Coyette, M., Graham, L.,
Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelis, C., Lacroque, K.,
Iamarares, R., Landers, T., Lehocaky, J., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKenna, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, J., Murphy, T., Naylor, J., Nguyen, C.,
Nicoli, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thoman, N.,
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (24-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 20, 2002 this sequence version replaced g1:21431101.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22881
Center clone name: 338_L_22
-----
Location/Qualifiers
1. 164107
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-338L22"
/clone_lib="RPC1-11 Human Male BAC"
complement (1..190)
/rpt_family="AluYo"
complement (245..426)
/rpt_family="Alu"
complement (427..738)
/rpt_family="AluSx"
complement (739..1453)
/rpt_family="Alu"
1454..1703
/rpt_family="TCCC)n"
unsure
1657..1660
/note="single clone coverage"
unsure
1665
/note="probably T, possibly TT"
complement (11704..1894)
/rpt_family="L1"
1907..2021
/rpt_family="L1M"
complement (2096..2233)
/rpt_family="L1M3A"
complement (2286..2511)
/rpt_family="L1M4"
3574..3747
/rpt_family="AluY"
3782..3824
/rpt_family="AT_rich"
3927..4011
/rpt_family="MIR"
4014..4265
/rpt_family="L2"
4747..5057
/rpt_family="AluSx"
5553..7200
/rpt_family="MER52A"
7354..7527
/rpt_family="MER5B"
7674..7791
/rpt_family="MIR"
complement (8537..8741)
/rpt_family="MIR"
8923..9081
/rpt_family="L1ME3A"
complement (9492..9805)
/rpt_family="AluYo"
complement (9811..9943)
/rpt_family="MER5B"
9944..10243
/rpt_family="AluSx"
complement (10244..10287)
/rpt_family="MER5B"
11025..11346
/rpt_family="AluSx"
12802..12858
/rpt_family="L2"
complement (13372..13631)
/rpt_family="AluYo"
complement (13813..14105)
/rpt_family="AluSx"
15481..15908
/rpt_family="MER6A"
16970..17272
/rpt_family="AluY"
17489..17628
/rpt_family="AluSx"
17629..17653
/rpt_family="T(A)n"
17654..17822
/rpt_family="AluSx"
18478..18780
/rpt_family="AluY"
18877..18898
/rpt_family="AT_rich"
19087..19398
/rpt_family="AluBq"
19411..19655
/rpt_family="AluYo"
19772..20060
/rpt_family="AluSx"
complement (20327..20621)
/rpt_family="AluY"
20742..20833
/rpt_family="L1M4"
complement (20859..21156)

```

repeat_region /rpt_family="AluSc"
21414..21583
/rpt_family="MIR"
repeat_region 21592..21739

Query Match 84.0%; Score 16.8; DB 9; Length 164107;
Best Local Similarity 90.0%; Pred. No.3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 119533 AGTCACAAACAGTACAGGG 119552

RESULT 32
AC083894/c 167273 bp DNA linear ROD 14-NOV-2002
LOCUS AC083894
DEFINITION Mus musculus chromosome 6 clone rp23-259j8 strain C57BL/6J,
complete sequence.
AC083894
VERSION AC083894.21 GI:24961434
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE 1 (bases 1 to 167273)
JOURNAL Unpublished
2 (bases 1 to 167273)
REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Nov 14, 2002 this sequence version replaced gi:22094402.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKXOR

FEATURES
source Location/Qualifiers

1..167273
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="6"

/clone="rp23-259j8"
/clone_id="RP01 - 23 Female (C57BL/6J) Mouse BAC Library"

Query Match 84.0%; Score 16.8; DB 10; Length 167273;
Best Local Similarity 90.0%; Pred. No.3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 5972 AGTTGAAACAGTACATGG 5953

RESULT 33
AC129943 178546 bp DNA linear HTG 06-AUG-2004
LOCUS AC129943
DEFINITION Mus musculus chromosome 5 clone RP23-410W21 map 5, *** SEQUENCING
IN PROGRESS ***.
AC129943
VERSION AC129943.7 GI:51011211
KEYWORDS HTG; HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
TITLE 1 (bases 1 to 178546)
JOURNAL Unpublished
2 (bases 1 to 178546)
REFERENCE
AUTHORS Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, J., Bouhassira, B.,
Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardya, S., Gird, S., Graham, L., Grand-pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, P., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retz, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178546)
REFERENCE
AUTHORS Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N.,
Anderson, M., Anderson, S., Atchachi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, J., Bouhassira, B., Camarata, J., Chang, J.,
Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galagan, J., Gardya, S., Graham, L., Grand-pierre, N., Hagos, B.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
McCarthy, M., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McLean, C., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, P., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retz, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKXOR

TITLE Direct Submission
JOURNAL Submitted (06-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2004 this sequence version replaced gi:50511617.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WITR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@road.mit.edu
Project Information
Center project name: 126468
Center clone name: 410_M21

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
1 178546: contig of 178546 bp in length.
Location/Qualifiers

1. 178546
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-410M21"
/clone_1ib="RP23-410M21"
/clone_1ib="RP23-410M21"
/clone_1ib="RP23-410M21"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 178546;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 157450 AGTCGCAACAGTACATGG 157469

RESULT 34
AC140318/c 180212 bp DNA linear HTG 14-JUL-2004
LOCUS Mus musculus chromosome 9 clone RP23-400F19, WORKING DRAFT
DEFINITION
SEQUENCE
AC140318
AC140318.3 GI:50284668
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEPIN.
KEYWORDS
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 180212)
Wilson, R.K.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 180212)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 180212)
Wilson, R.K.
Direct Submission
Submitted (14-JUL-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
On Jul 14, 2004 this sequence version replaced gi:29824550.

Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
Project Information
Center project name: M.BA0400F19

Summary Statistics

Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primed ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180137 bases at least Q40
Consensus quality: 180190 bases at least Q30
Consensus quality: 180203 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 180212: contig of 180212 bp in length.
Location/Qualifiers

1. 180212
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-400F19"
/clone="RP23-400F19"
/note="assembly_name:Contig24"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 180212;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 53960 AGTCGCAACAGTACATGG 53941

RESULT 35
AC147514 187725 bp DNA linear ROD 06-AUG-2004
LOCUS Mus musculus chromosome 1, clone RP23-442A6, complete sequence.
DEFINITION
SEQUENCE
AC147514
AC147514.11 GI:51011231
HTG.
KEYWORDS
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 187725)
Birren, B.; Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP23-442A6
Unpublished
2 (bases 1 to 187725)
Birren, B.; Nusbaum, C.; Lander, E.; Aboueleil, A.; Allen, N.; Anderson, M.; Arachchi, H.M.; Barna, N.; Bastien, V.; Bloom, T.; Boguslavsky, L.; Bonkagaiter, B.; Camarero, V.; Chang, V.; Choepel, Y.; Collymore, A.; Cook, A.; Cooke, P.; Corum, B.; DeRellano, K.; Diaz, V.S.; Dodge, S.; Dooley, K.; Dorris, L.; Erickson, J.; Fero, S.; Ferreira, P.; Fitzgerald, M.; Gage, D.; Galagan, J.; Gargana, S.; Graham, L.; Grand-Pierre, N.; Hafez, N.; Hagopian, D.; Hargis, B.; Hall, J.; Horton, L.; Hulme, W.; Iliev, I.; Johnson, R.; Jones, C.; Kamet, A.; Karatas, A.; Kells, C.; Landers, T.; Levine, R.;

TITLE
Submitted (07-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
(Dases 1 to 187725)

REFERENCE
3 (Dases 1 to 187725)

AUTHORS
Britten,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Menais,T., Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,V., Raymond,C., Retica,R., Rise,C., Rogov,P., Roman,J., Schaner,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbins,K., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zairoun,J., Zemke,L., Zimmer,A. and Zody,M.

```

repeat_region      20507..21000
                    /rpt_family="LxS"
repeat_region      21070..21104
                    /rpt_family="AT_rich"
repeat_region      21430..21563
                    /rpt_family="AT_rich"
repeat_region      22581..22651
                    /rpt_family="B1F"
repeat_region      22660..22667
                    /rpt_family="(A)n"
repeat_region      complement(22688..23503)
                    /rpt_family="RMR1A"
repeat_region      complement(23478..23989)
                    /rpt_family="RMR1A"
repeat_region      23992..24050
                    /rpt_family="(TTC)n"
repeat_region      complement(24055..24122)
                    /rpt_family="L1_MM"
repeat_region      24123..24146
                    /rpt_family="(CACCC)n"
repeat_region      complement(24147..24528)
                    /rpt_family="L1_MM"
repeat_region      24552..25808
                    /rpt_family="Lx"
repeat_region      25809..25912
                    /rpt_family="ORR1B-int"
repeat_region      25913..26317
                    /rpt_family="ORR1C"
repeat_region      26318..26399
                    /rpt_family="ORR1D"
repeat_region      27209..27235

```

Query Match 84.0%; Score 16.8; DB 10; Length 187725;

Best Local Similarity 90.0%; Pred. No. 3.1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
 |||||
 Db 69092 AGTCGCAACAGTACATGG 69111

RESULT 36
 AC079443 207091 bp DNA linear ROD 19-NOV-2003
 LOCUS Mus musculus clone rp23-381f7 map 6 strain C57BL/6J, complete

DEFINITION Mus musculus clone rp23-381f7 map 6 strain C57BL/6J, complete

ACCESSION AC079443
 VERSION AC079443
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.

TITLE Mus musculus BAC Clone rp23-381f7

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.

TITLE Direct Submission

AUTHORS Submitted (01-SEP-2000) Department Of Chemistry And Biochemistry,

JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

REFERENCE 3 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.

TITLE Direct Submission

AUTHORS Submitted (08-NOV-2003) Department Of Chemistry And Biochemistry,

JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 4 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.
 Direct Submission
 Submitted (18-NOV-2003) Department Of Chemistry And Biochemistry,

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

FEATURES
 source

Location/Qualifiers
 1..207091
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /map="6"
 /clone="rp23-381f7"
 /clone_id="RPCT - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 207091;

Best Local Similarity 90.0%; Pred. No. 3.1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
 |||||
 Db 19583 AGTCGCAACAGTACATGG 19602

RESULT 37

AC132022 229916 bp DNA linear HTG 15-NOV-2002

LOCUS Rattus norvegicus clone CH230-490L1, *** SEQUENCING IN PROGRESS

DEFINITION *** 5 unordered pieces.

ACCESSION AC132022

VERSION AC132022.3

KEYWORDS HTG, HTGS, PHASE1, HTGS_DRAFT, HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 229916)

Munzy,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Albrooks,S., Amiri,A., Angiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Behnmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryan,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Daya,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guetara,W.,

Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,

Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenshewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., McWhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankovics, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plodper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Remy, B., Remy, M., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Wolley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 229916)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (29-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229916)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:2398476.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: KBRB
Center clone name: CH230-490L11

----- Summary Statistics
Assembly program: Phrap; version: 0.990329
Consensus quality: 192103 bases at least Q40
Consensus quality: 195577 bases at least Q30
Consensus quality: 197895 bases at least Q20
Estimated insert size: 198652; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 23162: contig of 23162 bp in length
* 23163 23262: gap of unknown length
* 23263 190541: contig of 167279 bp in length
* 190542 190641: gap of unknown length
* 190642 203791: contig of 13150 bp in length
* 203792 203891: gap of unknown length
* 203892 222968: contig of 19077 bp in length
* 222969 223068: gap of unknown length
* 223069 223162: contig of 6848 bp in length.
Location/Qualifiers
1. .229916
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-490L11"
1. .3229
/note="wgs_end_extension
clone_end:Sp6"
6915. .8054
/note="clone boundary
clone_end:Sp6"
site:
end_sequence="B2128304"
23263. .24747
/note="wgs_contig"
complement(30107. .31256)
/note="clone boundary
clone_end:T7
site:
end_sequence="B2128302"
44736. .46308
/note="wgs_end_extension
clone_end:T7"
47013. .49058
/note="wgs_end_extension
clone_end:T7"
190642. .199974
/note="wgs_end_extension
clone_end:T7"
200388. .201787
/note="wgs_end_extension
clone_end:T7"
201835. .203791
/note="wgs_end_extension
clone_end:T7"
203892. .211452
/note="wgs_end_extension
clone_end:T7"
213252. .215045
/note="wgs_end_extension
clone_end:T7"
217112. .218575
/note="wgs_end_extension
clone_end:T7"
218820. .220272
/note="wgs_end_extension
clone_end:T7"
220774. .222968
/note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 229916;
Best Local Similarity 90.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATG 20
Db 142955 AGTCACAAACAGTACATG 142936

RESULT 38
AC112824
LOCUS
DEFINITION
AC112824 clone CH230-365H1, *** SEQUENCING IN PROGRESS

AC112824 GI:24635565
HTG: HTGS PHASE2: HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 240243)
Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amth,A, Anguiano,D, Anyadebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bismail,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flaggs,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Gervara,M, Gebregeorgis,E, Geer,K, Gill,K, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogue,M, Hollins,B, Howells,S, Huik,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,Z, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,D, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorensunwa,L, Loulsegad,H, Lozada,R,J, Lu,X, Ma,U, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Vailoy,K, Mangum,A, Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E, Manjoney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Mlosoajlevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L, Nankervs,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Parks,K, Nwackemele,O, Okwuonu,G, Olarunpunsagun,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Polindexter,A, Popovic,D, Primus,E, Pu,L, L, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reich,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Richard,S, Ruiz,S,J, Sanders,M, Saverly,G, Scherer,S, Scott,C, Shatsman,S, Sheen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J, Steimle,M, Strong,N, Thomas,S, Svattek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villanasa,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Woden,H, Morley,K, Williams,G, Wilson,R, Wleczky,R, Wooden,H, Morley,K, Wright,D, Wright,R, Wu,J, Yakub,X, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.
Direct Submission
Unpublished
2 (bases 1 to 240243)
Morley,K,C.
Direct Submission
Submitted (25-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
3 (bases 1 to 240243)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23603863.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRER
Center clone name: CH230-365H1
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 168357 bases at least Q40
Consensus quality: 170011 bases at least Q30
Consensus quality: 170942 bases at least Q20
Estimated insert size: 175728; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 240243: contig of 240243 bp in length.
Location/Qualifiers
1. 240243
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-365H1"
1. 1773
/note="wgs_contig"
ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 240243;
Best Local Similarity 90.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 AGTCGCAACAGTACATCG 20
193218 AGTCACACACGATGATG 193237
Db 193218
RESULT 39
AC120133
LOCUS
DEFINITION
IN PROGRESS
AC120133
AC120133 GI:39841417
282567 bp DNA linear HTG 15-DEC-2003
Mus musculus chromosome 6 clone RP23-112L19 map 6, *** SEQUENCING
Submitted (25-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
(bases 1 to 321695)
Rat Genome Sequencing Consortium.
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24819582.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be whole contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GHCM
Center clone name: CH230-51C23
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 261123 bases at least Q40
Consensus quality: 264929 bases at least Q30
Consensus quality: 267345 bases at least Q20
Estimated insert size: 281114; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces

is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 24448: contig of 24448 bp in length
* 24449: gap of unknown length
* 24549: contig of 18559 bp in length
* 43108: gap of unknown length
* 43208: contig of 4669 bp in length
* 47877: gap of unknown length
* 47977: gap of 264851 bp in length
* 312828: contig of unknown length
* 312927: gap of 1913 bp in length
* 312928: contig of 1913 bp in length
* 314841: gap of unknown length
* 314940: contig of 1499 bp in length
* 314941: gap of unknown length
* 316539: gap of 2094 bp in length
* 316540: contig of 2094 bp in length
* 318634: gap of unknown length
* 318634: contig of 2962 bp in length.
* 318734: Location/Qualifiers
1. .321695
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-51C23"
1. .1526
/note="wgs_end extension
clone end:5p6"
complement(6233. .6853)
/note="clone boundary
clone end:5p6
site:ECORI
end_sequence:BH337792"
15203. .24448
/note="wgs_contig"
43208. .44716
/note="wgs_contig"
45034. .47876
/note="wgs_contig"
47977. .49325
/note="wgs_contig"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 321695;
Best Local Similarity 90.0%; Pred. No.3e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

Qy 1 AGTCGCAACAGTACATG 20
Db 273918 AGTCACAAACAGTAGAATG 273937

RESULT 41
BX571870 348505 bp DNA linear BCT 26-SEP-2003
LOCUS BX571870
DEFINITION Photorhabdus luminescens subsp. laumondii T101 complete genome;
segment 12/17.
ACCESSION BX571870 BX470251
VERSION BX571870.1 GI:36786547
KEYWORDS complete genome.
SOURCE Photorhabdus luminescens subsp. laumondii T101
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photorhabdus.
REFERENCE
AUTHORS
1 Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Taourit S.,
Bocs S., Boursaux-Bude C., Chaudier M., Dassa E., Derose R.,
Derzelle S., Freysinet G., Gaudriault S., Givaudan A., Glaeser P.,
Medigue C., Lanois A., Powell K., Sigulier P., Wingate V.,
Zouine M., Boemare N., Danchin A. and Kunst F.
Complete genome sequence of the entomopathogenic bacterium

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Photobabds luminescens
Nat. Biotechnol. 11 (1) (2003) In press
Duchaud, F., Frangeul, L., Rusniok, C. and Kunst, F.
Direct Submission
Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:
lfrangeul@pasteur.fr, fkunst@pasteur.fr

FEATURES
source
location/Qualifiers

1. .348505

/organism="Photobabds luminescens subsp. laumondii T701"
/mol_type="genomic DNA"
/strain="T701"

/db_xref="taxon:243265"
complement (477. .3848)

/locus_tag="plu3231"

complement (477. .3848)

/locus_tag="plu3231"

/function="Unknown"
/note="unnamed protein product; Similar to putative
membrane protein of Y. pestis"

/codon_start=1

/transl_table=11

/protein_id="CAE15605.1"

/db_xref="GI:36786548"

/translation="MKVNRHGYGPGYIFITLSAGIFLMHIYNOQLLEERGLKLS
VWYAVATVPILILGLFGNAVMMHTRONKQKRLPERETADRSSQSLVSDGDR
MLRGRFMYKVRILLVVEVQVAVIAPGLTTMGVYRTLLMGSSLOAEPTA
QLMALRLRRYRPLNIAVATENQSAQVPMMDKALRMQAOOLRQAPYLMQVC
HSTMSOGEITOTVQCFPPRCPTDVAOTOLIVPLBERGMOLEKTDAPLCRL
SANTKOGGIAHNOQVTPWLAEGDVYSRGIMFSPLKQCTTRLPQIMPTNAGCV
LEBSRCFHGRVGMHTQTCRGLIVLGLMGAMLSFTTRDQAVQAIVDLKY
QENSDAQMALKDLNEVDRLQDRAHEGVYERFGLSQNKLLTALPDYAVNNR
LIRDKAAVYLAHLSALVNLPGSPPLARQAQGHQDQKAYLMNARPEKVAFLTR
LDNEBERAGIAPGVLLDTAMDLSFYENLAHPBWRITADKALIKORQLLOPG
KHVAEALYQOMQSVKYSIDLIRQMGTDAAALFTTSHVIGPMFTROMEGGIO
KILAKVASRKEALIDWVLSDNROAVASVSPBELKARLTARVETDPAVMDPINSIR
NAKTRLSVIDQILTMADGROSPLIAMNTAYOGETQOCPSSDSISAKTLIN
KTPRINIOQIGLHGMDETGPPLIANGENQAGSRLIADSTLNLQTFPLRVIVRL
KIQQNMSDDPPMANALAKTYEAGSTDLTDQYGRILASLGQWRSQGLTVE
PLAQAWQVYLPKATNLNSQMTALVDNMDRAFAGRYPFAGGSEISLPMGQFIRD
SGRIEPLTRQLGILHKAGNQMVEVNGQVQINPDLKALINLSQSLDLPADGS
OGRIEPLRARAVDVETDLTDGOKLRFNMESMOSPRMGETYKRGVWLTMVSN
AGARLFGDYQGMGLIRMLAKARLDESRYRLFTAPDGLPLTWILRTLGEPLA
LLKLRGFKLPXNIFVYMPGNHTTISAVDLIEE"

complement (3845. .4924)

/locus_tag="plu3232"

complement (3845. .4924)

/locus_tag="plu3232"

/function="Unknown"

/note="unnamed protein product; Highly similar to unknown
protein of Photobabds luminescens"

/codon_start=1

/transl_table=11

/protein_id="CAE15606.1"

/db_xref="GI:36786549"

/translation="MSPLRRLMPTIILLIVGLVYAVLKDTPDEYHAIVSSIPS
IPLVILFGAVLRYEYQCAAYAWEEASAEPRFOWQMSKCLAVTINILITPEQOC
FAPUGDAEIRIAPFPEQGRALSSDSRQLAKRDQPEQOCPEYRSLAQITIASD
DHASVESATYQWKISILPILNRDILDEDDDEGLIMVIAFCWSPENKKAYSB
FVSQALFSTKFAQEKLVLAGMGRPLPSJEGELIKDQILFEYNRIDKXNMHLM
TGVIVDLTELTINSHDSAFNLFTENSVALIDHTGPPGPPLSIFLASALTEALISL
SHIVINCLPESGASALYMTKEIHQ"

complement (3856. .3861)

/locus_tag="plu3231"

complement (3856. .3861)

/locus_tag="plu3231"

complement (3856. .3861)

/locus_tag="plu3231"

complement (4944. .5191)

/locus_tag="plu3233"

complement (4944. .5183)

/locus_tag="plu3233"

/function="Unknown"

/note="unnamed protein product; Hypothetical gene"

RBS

gene

CDS

/codon_start=1
/transl_table=11
/protein_id="CAE15607.1"
/db_xref="GI:36786550"
/translation="MIMFMNIGFSHTDQSRIVSYALFFYDKQKPVKIGSLIFL
AKQGMHMYIVDCCCPSPIDPLRISSEYVELAHT"
complement (5186. .5191)
/locus_tag="plu3233"
complement (5196. .5429)
/locus_tag="plu3234"
complement (5196. .5429)
/locus_tag="plu3234"
/function="Unknown"

/note="unnamed protein product; Similar to unknown protein
of Photobabds and some similarities with unknown
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15608.1"

/db_xref="GI:36786551"

/translation="MGVDMHIFELMQVLGIYLLLYTWPILIDVFTLVRIKKCI
KDKYGRASRLIPAVFMGSLIFIMINLVN"

complement (5517. .6623)

/locus_tag="plu3235"

complement (5517. .6623)

/locus_tag="plu3235"

/function="Unknown"

/note="unnamed protein product; Highly similar to unknown
protein of Photobabds"

/codon_start=1

/transl_table=11

/protein_id="CAE15609.1"

/db_xref="GI:36786552"

/translation="MNRKTLITLAVICTLMMVMPVYPIITLKEKVTENVMQ
PHCVGRILIDLPAGNVEAGIIFIDAGIETERYPPAPQORIRLREOELKMOVE
PKMPPLKAYRLQDNMGVIFPBNRSFGIPGPARVLAHLSNGVAFVTKMFLS
DDYKEDDPDNTYQGSREOYVLOTLEKMKELLSRISGRDTEIPTEGTICIPGF
ITGSHDEKEMTFRVYKSDSDRHTFVETLNDIQEHLEIRRGHSGFLSMKXITI
RKRRKELNGYIEVELTAPTEKIEEPETKIPQYKFLILANETIGDYKNPFARINLRN
DRSASPYSENELIKFWDVSTFRPRPAFRQ"

complement (6620. .8236)

/locus_tag="plu3236"

complement (6620. .8236)

/locus_tag="plu3236"

/function="Unknown"

/note="unnamed protein product; Similar to unknown
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15610.1"

/db_xref="GI:36786553"

/translation="MMSBPKENEIFIHPYDEGLPYVVPNARIENIVATCLKXAT
KIVPIFLPGWGSNLSKTEGSEVWRLNLSFDVGLWVCSASRYKTLIDNKITVD
DSGDITPDHTEKXKQFQCKQWGEJAHNSYGFELPMQAVLADDERLAEYCLAEGE
KTLRORMVMSLNAWGESEPLTEAEVDHSDYDLYPVHVGWYMLQSNESARLQYV
DKYLAFYGRKCATKVIILVTHSMGGLVAVYSSELNLRKXIGIYGVVWPAICAPPTY
KMKKGEEDVATGIVVYDGEDMTPVLAQSGPLDLPGEYRGMLHIDGWQTKPLP
KSPVPEIYVLEZDRWMLCETRFPLNDKDKKDKGASMSYKILRNHYPIEIEELS
KHAPNTVAYGSEKRLSYGVISWQASQDYNNKTEDYSGLTFDRPIYDPAVLEGAT
RMVQSVGSPQDILAKTKPLAPKPEGGGVVPSGRILAAAYLSLATEVDEHGAAY
KQNETNEIDFARLFTLRSIVKXVQAVKXGN"

complement (1018. .10762)

/locus_tag="plu3237"

complement (8237. .10762)

/locus_tag="plu3237"

complement (8237. .10762)

/locus_tag="plu3237"

/function="Unknown"

/note="unnamed protein product; Similar to YARG related
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15611.1"

/db_xref="GI:36786554"

/translation="MTNLTPTIIFDHSRYRLKVMKNKAPLDVLAFTGQALSGPRRYA

```
IEFTSAKGIIEPAOMIMOOASFTLSPAINPGIRMPPIPEPELRISVGVISGFCLLS
TSRDSRERVVTEPLRLALRSHOYAIVOKLSVPEIVKELLREHGFPGODPLESLAY
SYPRERQVWQYGEDDLRFQRLALRIGVITKTLADRLKIDIVEFYDORCDQFQVYL
PARNBSGHWGDEDAVVMGMEIAYQVEGKIATRDYDADARYPCGLANTBADYTRDRI
TYGEAYHVDYDRIIPGDARYAOPETESGVFARLHERYLNRILRLRMTTSATLVPG
CEIKVADAPETFRKCAITITQITNSARSDSEFMTFTAIPESEVCFPERVPRKVA
GITPAVSTKVDYDIDKDLRVFVFNFERKMPQGESLWRLARPYAGENVYF
HNPILGTEVAIAFEGGDDPRPIYALHDSVAPDHVNLVYKRVYLRTPAANKLMD
DRGRPHITLSTEGGKSQNLGHVDSORPHDPDRGGEFLRTDMDKAIARAGKLF
SHDKARACGEVILBMAALNQOQATLFEALYGAENAKAELADLQCKKALSTTLA
EAKSKALISAPBGIAQTAKSLQLAGSNITVTSKSTDFSVLKFVVAAGEBISLIF
AOKLGIKLPASKGRVDIEAGGDEMGLAALKDITVTSHEKVLISAKOELILVSGGYI
RIONGVCEGAPNHIIORAAVWQKEDGVSQTVQWQMTANYAVPKAVRAYKXISPLD
ROMQFHAEDGVOALSTIONGKTPLOKQGVIEISQLKIDEE"
complement(8242..8247)
/locus_tag="p1u3236"
complement(8242..8247)
/locus_tag="p1u3236"
complement(10767..10772)
/locus_tag="p1u3237"
complement(11601..14951)
/locus_tag="p1u3238"
complement(11601..14951)
/locus_tag="p1u3238"
/function="Unknown"
/Note="unnamed protein product; similar to putative
membrane protein of Y. pestis"
/codon_start=1
/transl_table=11
/protein_id="CAFI5612.1"
/db_xref="GI:36786555"
/translation="VVKTLILIGITLLIVSLAGLGMVWGEHRTTDAETFNHLIWC
GLAWVLVAVSMIGITVTLFVKNKVTDKSDTRPLPASEPDYDGDGDRRLVYGR
FMWYKRVILVGEVEGVAIAAGLTITMOGYSRTLLMGSLQAEPYVQALMLR
LRVYRPLNAILVATENQSAOPVMDKALRMOKOARQVLMQVAGSTMSOE
GRITQVGGFFPFCPTDIAVQLOLVNPLREMGQQLLEETAPLCRISANIKOO
GLAHNQVTLPTMAEYGVYSLRGMLFSLPLAQITRLPQMLPDTANGVLESGRCF
```

Query Match 84.0%; Score 16.8; DB 1; Length 348505;
Best Local Similarity 90.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTGCAACAGTACATGG 20
DB 11360 AGACGCAACAGTACAAAG 11379

RESULT 42
TAE566651 925 bp DNA linear PLN 16-MAR-2004
LOCUS Triticum aestivum subsp. aestivum partial Glu-A1-2 gene for
DEFINITION high-molecular-weight glutenin subunit, exon 1, specimen voucher
Gatersleben TRI 10914.
AJ566651
ACCESSION AJ566651.1 GI:32328626
VERSION Glu-A1-2 gene; high-molecular-weight glutenin subunit.
KEYWORDS Triticum aestivum (bread wheat)
SOURCE Triticum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE 1

Athlet, R.H., Jacomet, S. and Schlumberg, A.

Blatter, R.H., Jacomet, S. and Schlumberg, A.
About the origin of European spelt (Triticum spelta L.): allelic
differentiation of the HMW glutenin B1-1 and A1-2 subunit genes
Theor. Appl. Genet. 108 (2), 360-367 (2004)

JOURNAL 14564390

PUBMED 2 (bases 1 to 925)

REFERENCE Blatter, R.H.E.

AUTHORS Blatter, R.H.E.

TITLE Direct Submission

JOURNAL Submitted (18-MAY-2003) Blatter R.H.E., Genetics MPI for

Evolutionary Anthropology, Deutscher Platz, 04105 Leipzig, GERMANY

FEATURES Location/Qualifiers

```
source 1..925
/organism="Triticum aestivum"
/mol_type="genomic DNA"
/sub_species="aestivum"
/specimen_voucher="Gatersleben TRI 10914 (IPK)"
/db_xref="taxon:4565"
/tissue_type="leaf"
/country="Nepal"
/Note="cultivar-group: villosum"
423..925
/locus_tag="Glu-A1-2"
423..429
/locus_tag="Glu-A1-2"
514..>925
/locus_tag="Glu-A1-2"
/codon_start=1
/product="high-molecular-weight glutenin subunit"
/protein_id="GAB00513.1"
/db_xref="GI:32328627"
/db_xref="TrEMBL:Q7X6P9"
/translation="MAKRLVFAVTVIGLVSLTVABGEASQLOQCELEGGSLIACR
LVVDQLAGLPMSTGLQMRCCOQLRDISAKRPVALSVARQYGTAVPPKGSFVH
RETTQLQLOQGIQGTSSQTVGYRVSISPOG"
<514..>925
/locus_tag="Glu-A1-2"
/number=1
```

Query Match 82.0%; Score 16.4; DB 8; Length 925;
Best Local Similarity 94.4%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGCAACAGTACATGG 20
DB 874 TCGCAACAGTACAAAG 891

RESULT 43
CQ574602 950 bp DNA linear PAT 02-FEB-2004
LOCUS CQ574602
DEFINITION Sequence 2360 from Patent WO0171042.
ACCESSION CQ574602
VERSION CQ574602.1 GI:41638381
KEYWORDS Drosophila sp.
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1

Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.

Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses therefor
Patent: WO 0171042-A 2360 27-SEP-2001;

PE Corporation (NY) (US)

Location/Qualifiers

1..950

/organism="Drosophila sp."

/mol_type="unassigned DNA"

/db_xref="taxon:7242"

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 950;

Best Local Similarity 94.4%; Pred. No. 6.6e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGCAACAGTACATGG 20

DB 634 TCGCAACAGTACATGG 651

RESULT 44

AY060813

LOCUS AY060813 1143 bp mRNA linear INV 08-NOV-2001
DEFINITION Drosophila melanogaster GH28342 full length cDNA.
ACCESSION AY060813
VERSION AY060813.1 GI:16768283
KEYWORDS F11 cDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
AUTHORS Neupert, M.; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Eukaryota; Metazoa; Arthropoda; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1143)
Stapleton, M., Broks, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nuno, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celiker, S.
Direct Submission
Submitted (30-OCT-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription errors that result in single base changes, reverse transcription errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to conae@fruitfly.berkeley.edu.
FEATURES
source
1..1143
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="46A3-46A3"
1..1143
/gene="CG1773"
/note="alignment with genomic scaffold AE003832"
/db_xref="FLYBASE:FBgn0033439"
9..962
/gene="CG1773"
/note="Longest ORF"
/codon_start=1
/product="GH28342p"
/protein_id="AAL28361.1"
/db_xref="GI:16768284"
/translation="MNSPLGIRALIWGILICSCPPSSQAGREDWPHETLAVEDTQDGVLSNLIIPAGRLRRITGGRKSSLSPPMAFTHISGDIEMCRGSSLSSELVITAAHCFMCPKREIRWVIGELDISSTDCVINYRVALPYEETIDKWLHEFNLFYEGDIALIKLNKRVFDMHRIPLCLPTDELLAFTLQGSYAAVGMGRTERRFA NSTMEVHINTEKCTDGRDTSFLCANGDYVDCTGSGSLPWKTLIFKARTVOFGVYSTGSGCAGQAYVMDVPTVYFMIILAKVAELSDPKSLHR"
ORIGIN
Query Match 82.0%; Score 16.4; DB 3; Length 1143;
Best Local Similarity 94.4%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACATGG 20
|||||
Db 658 TCGCCACAGTACATGG 675

RESULT 45
LOCUS CQ594254/c 2901 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 22012 from Patent WO0171042.
ACCESSION CQ594254
VERSION CQ594254.1 GI:1651482
KEYWORDS
SOURCE
ORGANISM Drosophila sp.
Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
Venter, J. C., Adams, M., Li, P. W. and Myers, E. W.
Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof
Patent: WO 0171042-A 22012 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source
1..2901
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Query Match 82.0%; Score 16.4; DB 6; Length 2901;
Best Local Similarity 94.4%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACATGG 20
|||||
Db 2408 TCGCCACAGTACATGG 2391
|||||
RESULT 46
LOCUS CQ574601/c 3165 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 2359 from Patent WO0171042.
ACCESSION CQ574601
VERSION CQ574601.1 GI:1638380
KEYWORDS
SOURCE
ORGANISM Drosophila sp.
Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
Venter, J. C., Adams, M., Li, P. W. and Myers, E. W.
Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof
Patent: WO 0171042-A 2359 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source
1..3165
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Query Match 82.0%; Score 16.4; DB 6; Length 3165;
Best Local Similarity 94.4%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACATGG 20
|||||
Db 1317 TCGCCACAGTACATGG 1300
|||||
RESULT 47
LOCUS AC017970 89765 bp DNA linear HTG 09-DEC-1999
AC017970

DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.

ACCESSION AC01970

VERSION AC01970.1 GI:6553220

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

REFERENCE Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS 1 (bases 1 to 89765)

TITLE Adams, M. and Venter, J.C.

JOURNAL Direct Submission

COMMENT Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

FEATURES

source

1..89765

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

ORIGIN

Query Match 82.0%; Score 16.4; DB 2; Length 89765;

Best Local Similarity 94.4%; Pred. No. 5.3e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCGCAACAGTACATCG 20

|||||

Db 81121 TCGCAACAGTACATCG 81138

RESULT 48

AL390057 100578 bp DNA linear PRI 25-MAR-2001

LOCUS Human DNA sequence from clone RP11-26118 on chromosome 6 Contains

DEFINITION STS and GSSs, complete sequence.

ACCESSION AL390057

VERSION AL390057.12 GI:11322100

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 100578)

AUTHORS Williams, S.

TITLE Direct Submission

JOURNAL Submitted (23-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, UK

COMMENT Submitted (23-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, UK

CH10 15A, UK. E-mail enquiries: humquerry@sanger.ac.uk

On Nov 23, 2000 this sequence version replaced gi:1139956.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/chr6>

IMPORTANT: This sequence is not the entire insert of clone RP11-26118 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

FEATURES source

The true right end of clone RP11-26118 is at 100578 in this sequence. The true left end of clone RP11-300612 is at 62151 in this sequence. The true right end of clone RP11-67821 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-26118 is from the library RP11-1.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/Dacpac/home.htm>

VECTOR: PBAC3.6.

Location/Qualifiers

1..100578

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/clone_id="RP11-26118"

/clone_id="RP11-1.1"

101..340

/note="match: GSS: Em:AQ302222"

155..1072

/note="U1PA3 repeat: matches 5230..6146 of consensus"

1078..1291

/note="U1PA3 repeat: matches 5022..5235 of consensus"

complement(3326..3431)

/note="match: GSS: Em:AQ043864"

6539..6945

/note="match: GSS: Em:AQ296519"

9258..10681

/note="U1PA7 repeat: matches 3942..5399 of consensus"

10984..11725

/note="U1PA7 repeat: matches 5399..6124 of consensus"

complement(112748..13028)

/note="match: GSS: Em:AQ618053"

12918..13165

/note="match: GSS: Em:AQ054828"

12990..13356

/note="match: GSS: Em:AQ588764"

13504..13645

/note="U1Mcb repeat: matches 789..930 of consensus"

13901..13940

/note="U1Mcb repeat: matches 789..930 of consensus"

14079..14422

/note="U1Mcb repeat: matches 143..443 of consensus"

15727..15784

/note="U1Mcb repeat: matches 143..443 of consensus"

16035..16272

/note="U1Mcb repeat: matches 6239..6296 of consensus"

16035..16272

/note="match: GSS: Em:AQ042985"

20715..20810

/note="U1Mcb repeat: matches 6098..6304 of consensus"

22827..23016

/note="U1Mcb repeat: matches 6098..6304 of consensus"

complement(125458..26180)

/note="match: GSS: Em:AQ486821"

26590..28301

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

30142..30175

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

30620..30675

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

30884..32217

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

34613..34688

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

34773..36379

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

36240..36273

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

36380..36413

/note="U1Mcb internal repeat: matches 3..1579 of consensus"

```

repeat_region /note="17 copies 2 mer aa 82% conserved"
/note="LIPB5 repeat: matches 5770. .6140 of consensus"
38583.38816
/note="LIM4 repeat: matches 5931. .6164 of consensus"
41526. .41977
/note="match: GSS: Em:AQ126191"
41540. .42000
/note="match: GSS: Em:AQ613376"
41548. .42098
/note="match: GSS: Em:AQ378516"
41552. .42000
/note="match: GSS: Em:AQ703614"
42853. .43035
/note="LIP1A4 repeat: matches 5962. .6144 of consensus"
43625. .43704
/note="40 copies 2 mer ag 73% conserved"
47701. .48062
/note="LIP1A2 repeat: matches 3. .374 of consensus"
complement(48868. .49490)
/note="match: GSS: Em:AQ109429"
49283. .49723
/note="match: GSS: Em:AQ803593"
49524. .49555
/note="16 copies 2 mer tg 90% conserved"
50381. .50723
/note="TIGER1 repeat: matches 2045. .2415 of consensus"
51284. .51307
/note="12 copies 2 mer aa 100% conserved"
56705. .57064
/note="THIC repeat: matches 1. .371 of consensus"
57065. .58634 repeat: matches 1. .1580 of
/note="THIC-internal repeat: matches 1. .1580 of
consensus"
58637. .58939
/note="THIC repeat: matches 5. .371 of consensus"
60208. .60748
/note="match: GSS: Em:AQ389684"
60770. .60825
/note="28 copies 2 mer aa 71% conserved"
complement(60839. .61518)
/note="match: GSS: Em:AQ308890"
62037. .62244
/note="LIM4 repeat: matches 6102. .6295 of consensus"
62539. .62734
/note="LIM4 repeat: matches 5917. .6115 of consensus"
62854. .65962
/note="LIM4 repeat: matches 2760. .5917 of consensus"
65369. .68471
/note="LIP3 repeat: matches 3648. .6146 of consensus"
68476. .72115
/note="LIP3 repeat: matches 11. .3642 of consensus"
72327. .73908
/note="LIP3 repeat: matches 4581. .6148 of consensus"
75859. .76732
/note="MER1D repeat: matches 1. .897 of consensus"
76897. .77556
/note="PAB1A repeat: matches 1. .660 of consensus"
79727. .79864
/note="68 copies 2 mer aa 68% conserved"
83250. .83780
/note="match: GSS: Em:AQ566452"
83271. .83732
/note="match: GSS: Em:AQ420577"
complement(85140. .85328)
/note="match: GSS: Em:HSP331F2"
complement(86967. .89333)
/note="match: GSS: Em:AQ242452"
89334. .89573
/note="MER30 repeat: matches 1. .230 of consensus"
92101. .92396
/note="match: GSS: Em:AQ66482"
93663. .93702

```

```

repeat_region /note="20 copies 2 mer ta 82% conserved"
95778. .97415
/note="LIPB2 repeat: matches 4513. .6155 of consensus"
97414. .98915
/note="LIPB2 repeat: matches 2970. .4484 of consensus"
repeat_region /note="139 copies 2 mer ta 71% conserved"
99680. .99757
/note="complement(99955. .100244)
complement(99955. .100244)
/note="match: GSS: Em:AQ388863"
complement(99960. .100368)
/note="match: GSS: Em:AQ388862
match: STS: Em:G59389"
complement(100039. .100572)
/note="match: GSS: Em:AQ416636"
complement(100159. .100554)
/note="match: GSS: Em:AQ003409
match: STS: Em:G51488"

ORIGIN
Query Match 82.0%; Score 16.4; DB 9; Length 100578;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 GTGGCAACAGTCAATG 19
Db 80450 GTACCAACAGTCAATG 80467

RESULT 49
BX649472 108623 bp DNA linear VRT 17-FEB-2004
LOCUS Zebrafish DNA sequence from clone CH211-214K in linkage group 20,
DEFINITION complete sequence.
ACCESSION BX649472
VERSION BX649472.15 GI:42592727
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 108623)
REFERENCE
AUTHORS Dyer, L.
TITLE Direct Submission
JOURNALS Submitted (14-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 17, 2004 this sequence version replaced GI:41222936.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., paired quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; InfomRNA,
on the WormPep database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived

```


zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-214K9 is from a CHOR1-211 BAC library

FEATURES

source

location/Qualifiers
1..108623
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-214K9"
/clone_lib="CHOR1-211"

ORIGIN

Query Match 82.0%; Score 16.4; DB 5; Length 108623;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAT 18
|||
Db 46618 AGTTCGCAACAGTACAT 46635

RESULT 50
AL365229/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-723K16 on chromosome 20 Contains ESTs, STSs, GSSs and two Cpg islands. Contains the 5' end of the CDH4 gene for retinal cadherin 4 type 1 (R-cadherin), complete sequence.

ACCESSION AL365229 121478 bp DNA linear PRI 14-FEB-2001
VERSION AL365229.GI:10443454
KEYWORDS HTG; cadherin; CDH4; Cpg island.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 121478)
TITLE Wilson, S.
JOURNAL Direct Submission

Submitted (09-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Oct 1, 2000 this sequence version replaced gi:10186767.
During difference assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMBPP; Information on the WORMBPP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpp This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP11-723K16. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true right end of clone RP11-723K16 is at 121479 in this sequence. The true right end of clone RPS-1040G13 is at 100 in this sequence. RP11-723K16 is from the library RPCI-11.3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

source

VECTOR: pBAC3.6
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

location/Qualifiers

1..121478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-723K16"
/clone_lib="RPCI-11.3"
complement(1..76)
/note="match: STS: Em:HS040G13S"
1480..1588
/note="12 repeat: matches 2606..2716 of consensus"
1580..1920
/note="match: GSS: Em:BL7526"
1612..1738
/note="match: GSS: Em:BL5710"
1683..2095
/note="MER66A repeat: matches 27..471 of consensus"
2356..2603
/note="ULMC4 repeat: matches 6694..6947 of consensus"
3193..3283
/note="WIR repeat: matches 118..221 of consensus"
3619..3865
/note="WLR1H repeat: matches 293..547 of consensus"
3866..4159
/note="ALUSx repeat: matches 1..293 of consensus"
4160..4314
/note="WLR1H repeat: matches 1..293 of consensus"
4894..4929
/note="9 copies 4 mer ttta 83% conserved"
7552..7681
/note="MER5B repeat: matches 38..166 of consensus"
7906..8195
/note="ALUSx repeat: matches 1..296 of consensus"
8365..8779
/note="ULMC3 repeat: matches 1245..1663 of consensus"
9548..9583
/note="18 copies 2 mer tg 88% conserved"
9554..9585
/note="8 copies 4 mer tgtg 96% conserved"
10035..11189
/note="ULMC4 repeat: matches 6696..7851 of consensus"
11976..12219
/note="WLR1J repeat: matches 275..514 of consensus"
12442..12502
/note="ULMC3 repeat: matches 6042..6102 of consensus"
12552..12746
/note="3 copies 65 mer 81% conserved"
12556..12609
/note="9 copies 6 mer tcgag 77% conserved"
12615..12680
/note="11 copies 6 mer tgltag 83% conserved"
12685..13044
/note="12 copies 30 mer 77% conserved"
12700..13053
/note="59 copies 6 mer tagtg 78% conserved"
12856..12859
/note="Single clone region"
13050..13289
/note="16 copies 15 mer 64% conserved"
13053..13256
/note="68 copies 3 mer gtc 65% conserved"
13072..13311
/note="8 copies 30 mer 87% conserved"
13098..13299

```
/note="101 copies 2 mer gt 55% conserved"
repeat_region 13253..13382
/note="2 copies 65 mer 83% conserved"
repeat_region 13255..13320
/note="11 copies 6 mer tctagg 75% conserved"
repeat_region 13316..13687
/note="124 copies 3 mer gag 55% conserved"
repeat_region 13321..13686
/note="61 copies 6 mer gtagct 78% conserved"
repeat_region 13687..13728
/note="7 copies 6 mer gtgcag 81% conserved"
repeat_region 13398..14489
/note="MER53 repeat: matches 1..95 of consensus"
repeat_region 14502..14794
/note="AluX repeat: matches 1..233 of consensus"
repeat_region 15370..15476
/note="L2 repeat: matches 2625..2742 of consensus"
repeat_region 16777..17289
/note="MLTID repeat: matches 2..505 of consensus"
repeat_region 17603..17729
/note="MIR repeat: matches 28..153 of consensus"
repeat_region 17810..17887
/note="MAD1 repeat: matches 1..80 of consensus"
repeat_region 18309..18503
/note="MER3 repeat: matches 11..209 of consensus"
repeat_region 19972..20149
/note="MER5B repeat: matches 1..177 of consensus"
repeat_region 20569..20977
/note="LIPB1 repeat: matches 5747..6152 of consensus"
repeat_region 21167..21517
/note="LIMB7 repeat: matches 5769..6130 of consensus"
repeat_region 22885..23003
/note="MLTID repeat: matches 1..121 of consensus"
repeat_region 23015..23157
/note="MLTID repeat: matches 261..452 of consensus"
repeat_region 23618..23673
/note="2 copies 28 mer 92% conserved"
repeat_region 23619..23674
/note="8 copies 7 mer ctgggct 92% conserved"
repeat_region 24512..24579
/note="17 copies 4 mer atat 100% conserved"
repeat_region 24632..24934
/note="AluX repeat: matches 1..306 of consensus"
misc_feature complement(24989..25460)
/note="match: GSS: Em:AQ721413"
repeat_region 25628..26093
/note="LIMC/D repeat: matches 5331..5763 of consensus"
repeat_region 26125..26179
/note="LIMB1 repeat: matches 6100..6156 of consensus"
repeat_region 26839..27148
/note="MLTIA2 repeat: matches 1..316 of consensus"
repeat_region 27569..27718
/note="3 copies 50 mer 78% conserved"
repeat_region 27570..27719
/note="5 copies 30 mer 74% conserved"
repeat_region 27597..27716
/note="8 copies 15 mer 65% conserved"
repeat_region 27908..28211
/note="LIM3 repeat: matches 216..184 of consensus"
repeat_region 28586..28685
/note="25 copies 4 mer tctg 74% conserved"
repeat_region 28591..28678
/note="4 copies 22 mer 76% conserved"
repeat_region 28592..28675
/note="3 copies 28 mer 81% conserved"
repeat_region 28681..28830
/note="5 copies 30 mer 73% conserved"
repeat_region 28687..28818
/note="22 copies 6 mer tctgctg 77% conserved"
repeat_region 28688..28819
/note="3 copies 44 mer 83% conserved"
repeat_region 28693..28820
/note="32 copies 4 mer tctg 78% conserved"
```

```
repeat_region 28703..28812
/note="5 copies 22 mer 80% conserved"
misc_feature complement(30513..30994)
/note="match: GSS: Em:AQ454650"
repeat_region 30702..30783
/note="2 copies 41 mer 87% conserved"
repeat_region 31320..31610
/note="AluX repeat: matches 4..294 of consensus"
repeat_region 32334..32778
/note="L1 repeat: matches 2672..3123 of consensus"
repeat_region 32763..33175
/note="LIMB1 repeat: matches 5675..6089 of consensus"

Query Match 82.0%; Score 16.4; DB 9; Length 121478;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 AGTGCACACAGTACAT 18
Db 5869 AGTGCACACAGTACAT 5852
```

Search completed: December 3, 2004, 03:07:41
Job time : 704.579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 02:05:35 ; Search time 1585.42 Seconds
(without alignments)
62.385 Million cell updates/sec

Title: US-10-050-189a-6
Perfect score: 18
Sequence: 1 gagacaacaagatcgc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3694831 segs, 274740616 residues
Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	US-10-050-189a-6	Sequence 6, Appl1
2	16	88.9	2465	US-10-027-633-111967	Sequence 111967,
3	16	88.9	2465	US-10-027-633-111967	Sequence 111967,
4	16	88.9	66479	US-10-041-856-1	Sequence 1, Appl1
5	15.4	85.6	409	US-09-833-381-908	Sequence 908, Appl1
6	15.4	85.6	482	US-09-878-178-1155	Sequence 1155, Ap
7	15.4	85.6	482	US-10-046-935-1155	Sequence 1155, Ap
8	15.4	85.6	482	US-10-146-502-1155	Sequence 1155, Ap
9	15.4	85.6	513	US-10-282-122A-20772	Sequence 20772, A
10	15.4	85.6	577	US-10-027-632-277903	Sequence 277903, A
11	15.4	85.6	577	US-10-027-632-277903	Sequence 277903, A
12	15.4	85.6	582	US-10-027-632-277719	Sequence 277719, A

13	15.4	85.6	582	US-10-027-632-277719	Sequence 277719,
14	15.4	85.6	582	US-10-767-701-3256	Sequence 3256, Ap
15	15.4	85.6	775	US-10-437-963-32919	Sequence 32919, A
16	15.4	85.6	1470	US-09-815-242-6118	Sequence 6118, Ap
17	15.4	85.6	1470	US-10-369-493-32485	Sequence 32485, A
18	15.4	85.6	1518	US-10-369-493-32483	Sequence 32483, A
19	15.4	85.6	1950	US-10-369-493-45765	Sequence 45765, A
20	15.4	85.6	2186	US-10-169-257-8	Sequence 8, Appl1
21	15.4	85.6	3178	US-09-308-203-7	Sequence 7, Appl1
22	15.4	85.6	5101	US-10-093-463-135	Sequence 135, App
23	15.4	85.6	5719	US-10-367-094-100	Sequence 100, App
24	15.4	85.6	7043	US-09-070-927A-288	Sequence 288, App
25	15.4	85.6	142976	US-10-367-094-99	Sequence 99, Appl
26	15	83.3	378	US-10-437-963-86110	Sequence 86110, A
27	15	83.3	485	US-10-137-036-4	Sequence 4, Appl1
28	15	83.3	737	US-09-770-149-77	Sequence 77, Appl
29	15	83.3	1169	US-10-767-701-11437	Sequence 11437, A
30	15	83.3	1455	US-10-032-585-6133	Sequence 6133, Ap
31	14.8	82.2	209	US-10-425-115-87011	Sequence 87011, A
32	14.8	82.2	352	US-09-867-550-15975	Sequence 15975, Ap
33	14.8	82.2	416	US-10-425-115-91075	Sequence 91075, A
34	14.8	82.2	438	US-10-282-122A-33603	Sequence 33603, A
35	14.8	82.2	452	US-10-425-114-3291	Sequence 3291, Ap
36	14.8	82.2	462	US-10-767-701-26974	Sequence 26974, A
37	14.8	82.2	465	US-09-918-995-1845	Sequence 1845, Ap
38	14.8	82.2	467	US-10-360-238-4785	Sequence 4785, Ap
39	14.8	82.2	470	US-10-425-115-81576	Sequence 81576, A
40	14.8	82.2	473	US-10-027-632-68850	Sequence 68850, A
41	14.8	82.2	473	US-10-027-632-68851	Sequence 68851, A
42	14.8	82.2	473	US-10-027-632-68850	Sequence 68850, A
43	14.8	82.2	473	US-10-027-632-68851	Sequence 68851, A
44	14.8	82.2	480	US-10-424-599-16423	Sequence 16423, A
45	14.8	82.2	489	US-10-424-599-85425	Sequence 85425, A
46	14.8	82.2	585	US-10-027-632-217259	Sequence 217259, A
47	14.8	82.2	585	US-10-027-632-217260	Sequence 217260, A
48	14.8	82.2	585	US-10-027-632-217259	Sequence 217259, A
49	14.8	82.2	585	US-10-027-632-217260	Sequence 217260, A
50	14.8	82.2	666	US-10-425-114-5596	Sequence 5596, Ap
51	14.8	82.2	671	US-10-027-632-300982	Sequence 300982, A
52	14.8	82.2	671	US-10-027-632-300982	Sequence 300982, A
53	14.8	82.2	676	US-10-313-542-15	Sequence 15, Appl
54	14.8	82.2	689	US-10-425-114-16588	Sequence 16588, A
55	14.8	82.2	699	US-10-425-114-7195	Sequence 7195, Ap
56	14.8	82.2	701	US-10-425-115-115987	Sequence 115987, A
57	14.8	82.2	723	US-10-437-963-16634	Sequence 16634, A
58	14.8	82.2	727	US-10-425-114-23906	Sequence 23906, A
59	14.8	82.2	733	US-10-425-114-7249	Sequence 7249, Ap
60	14.8	82.2	736	US-10-425-114-1610	Sequence 1610, Ap
61	14.8	82.2	736	US-10-425-114-22570	Sequence 22570, A
62	14.8	82.2	741	US-10-425-114-18307	Sequence 18307, A
63	14.8	82.2	742	US-10-425-114-25243	Sequence 25243, A
64	14.8	82.2	765	US-10-425-114-1374	Sequence 1374, Ap
65	14.8	82.2	784	US-10-767-795-3053	Sequence 3053, Ap
66	14.8	82.2	795	US-10-653-047-4852	Sequence 4852, Ap
67	14.8	82.2	818	US-10-425-114-21956	Sequence 21956, A
68	14.8	82.2	857	US-10-425-114-21924	Sequence 21924, A
69	14.8	82.2	857	US-10-425-115-5872	Sequence 5872, Ap
70	14.8	82.2	858	US-10-282-122A-41774	Sequence 41774, A
71	14.8	82.2	900	US-10-437-963-73718	Sequence 73718, A
72	14.8	82.2	931	US-10-425-115-114244	Sequence 114244, A
73	14.8	82.2	943	US-10-363-454A-25013	Sequence 25013, A
74	14.8	82.2	943	US-10-363-454A-25014	Sequence 25014, A
75	14.8	82.2	944	US-10-767-701-13612	Sequence 13612, A
76	14.8	82.2	975	US-10-767-701-10956	Sequence 10956, A
77	14.8	82.2	1014	US-10-653-047-575	Sequence 575, App
78	14.8	82.2	1050	US-10-437-963-15092	Sequence 15059, A
79	14.8	82.2	1167	US-10-437-963-96110	Sequence 96110, A
80	14.8	82.2	1212	US-10-250-238-796	Sequence 796, App
81	14.8	82.2	1245	US-10-437-963-70738	Sequence 70738, A
82	14.8	82.2	1315	US-10-369-493-32991	Sequence 29991, A
83	14.8	82.2	1326	US-10-369-493-329042	Sequence 43042, A
84	14.8	82.2	1381	US-10-767-701-13509	Sequence 13509, A
85	14.8	82.2	1383	US-10-282-122A-41827	Sequence 41827, A

C 86	14.8	82.2	1412	17	US-10-437-963-6447	Sequence 6447, App	159	14.4	80.0	1168	18	US-10-425-115-177042	Sequence 177042,
87	14.8	82.2	1511	18	US-10-739-930-2432	Sequence 2432, App	C 160	14.4	80.0	1200	17	US-09-985-153-23	Sequence 23, Appl
88	14.8	82.2	1624	16	US-10-425-114-25941	Sequence 25941, A	161	14.4	80.0	1212	11	US-10-437-963-12980	Sequence 12980, A
89	14.8	82.2	1840	16	US-10-425-114-9032	Sequence 9032, Ap	162	14.4	80.0	1262	17	US-10-437-963-62663	Sequence 62663, A
90	14.8	82.2	1859	16	US-10-424-599-108745	Sequence 108745, A	163	14.4	80.0	1416	16	US-10-425-114-33433	Sequence 33433, A
91	14.8	82.2	1867	13	US-10-027-633-258922	Sequence 258922, A	164	14.4	80.0	1423	9	US-09-764-877-3760	Sequence 3760, Ap
92	14.8	82.2	1867	13	US-10-027-633-258922	Sequence 258922, A	165	14.4	80.0	1423	6	US-10-424-515-3760	Sequence 3760, Ap
93	14.8	82.2	1867	15	US-10-027-633-258922	Sequence 258922, A	166	14.4	80.0	1460	16	US-10-425-114-30890	Sequence 30890, A
94	14.8	82.2	1867	15	US-10-027-633-258922	Sequence 258922, A	C 167	14.4	80.0	1462	8	US-08-781-986A-771	Sequence 771, App
95	14.8	82.2	1923	18	US-10-425-115-19404	Sequence 19404, A	C 168	14.4	80.0	1462	16	US-10-329-624-771	Sequence 771, App
C 96	14.8	82.2	1954	16	US-10-425-114-26365	Sequence 26365, A	169	14.4	80.0	1662	17	US-10-437-963-62661	Sequence 62661, A
C 97	14.8	82.2	2014	16	US-10-425-114-12937	Sequence 12937, A	170	14.4	80.0	1823	15	US-10-037-270-516	Sequence 516, App
C 98	14.8	82.2	2084	16	US-10-425-114-27735	Sequence 27735, A	171	14.4	80.0	1823	15	US-10-117-722-516	Sequence 516, App
C 99	14.8	82.2	2174	18	US-10-425-115-33959	Sequence 33959, A	172	14.4	80.0	1834	16	US-10-425-114-4449	Sequence 4449, App
C 100	14.8	82.2	2174	18	US-10-425-115-33959	Sequence 33959, A	C 173	14.4	80.0	1890	16	US-10-425-115-19682	Sequence 19682, A
C 101	14.8	82.2	2468	16	US-10-282-122A-26301	Sequence 26301, A	174	14.4	80.0	1906	16	US-10-425-115-71755	Sequence 71755, A
C 102	14.8	82.2	2495	17	US-10-437-963-74259	Sequence 74259, A	175	14.4	80.0	1920	18	US-10-425-115-171486	Sequence 171486, A
C 103	14.8	82.2	2542	18	US-10-425-115-115985	Sequence 115985, A	176	14.4	80.0	1968	17	US-10-437-963-18069	Sequence 18069, A
C 104	14.8	82.2	2701	17	US-10-437-963-36790	Sequence 36790, A	C 177	14.4	80.0	2157	18	US-10-425-115-127123	Sequence 127123, A
C 105	14.8	82.2	2712	16	US-10-282-122A-8910	Sequence 8910, Ap	178	14.4	80.0	2369	15	US-10-104-047-927	Sequence 927, Appl
C 106	14.8	82.2	2809	18	US-10-425-115-146651	Sequence 146651, A	179	14.4	80.0	2390	14	US-10-149-819-23	Sequence 819, A
C 107	14.8	82.2	2821	15	US-10-278-536-551	Sequence 536, App	C 180	14.4	80.0	2400	16	US-10-424-559-23999	Sequence 23999, A
C 108	14.8	82.2	2821	15	US-10-225-068A-29	Sequence 51, Appl	181	14.4	80.0	2465	16	US-10-425-115-55899	Sequence 55899, App
C 109	14.8	82.2	2821	16	US-10-374-780A-77	Sequence 77, Appl	182	14.4	80.0	2670	17	US-10-437-963-12571	Sequence 12571, A
C 110	14.8	82.2	2821	16	US-10-412-699B-283	Sequence 283, App	183	14.4	80.0	2961	16	US-10-425-114-34064	Sequence 34064, A
C 111	14.8	82.2	2821	16	US-10-412-699B-283	Sequence 283, App	184	14.4	80.0	3125	16	US-10-264-049-868	Sequence 868, App
C 112	14.8	82.2	2831	16	US-10-425-115-114245	Sequence 114245, A	185	14.4	80.0	3296	9	US-09-909-320-310	Sequence 310, App
C 113	14.8	82.2	3554	15	US-10-369-493-37050	Sequence 37050, A	186	14.4	80.0	3296	9	US-09-909-088B-310	Sequence 310, App
C 114	14.8	82.2	4854	16	US-10-282-122A-3214	Sequence 32184, A	187	14.4	80.0	3296	9	US-09-905-291A-310	Sequence 310, App
C 115	14.8	82.2	6245	17	US-10-437-963-45500	Sequence 45500, A	188	14.4	80.0	3296	9	US-09-902-853-310	Sequence 310, App
C 116	14.8	82.2	40000	17	US-10-437-963-45500	Sequence 45500, A	189	14.4	80.0	3296	9	US-09-907-824-310	Sequence 310, App
C 117	14.8	82.2	60057	13	US-10-087-192-700	Sequence 700, App	190	14.4	80.0	3296	9	US-09-907-841-310	Sequence 310, App
C 118	14.8	82.2	84048	16	US-10-029-148B-1	Sequence 1, Appl	191	14.4	80.0	3296	9	US-09-904-011-310	Sequence 310, App
C 119	14.8	82.2	248436	13	US-10-087-192-2014	Sequence 2014, Ap	192	14.4	80.0	3296	10	US-09-903-640-310	Sequence 310, App
C 120	14.8	82.2	352938	17	US-10-322-656-79	Sequence 79, Appl	193	14.4	80.0	3296	10	US-09-908-093-310	Sequence 310, App
C 121	14.8	82.2	430442	18	US-10-417-375-128	Sequence 128, App	194	14.4	80.0	3296	10	US-09-905-747A-310	Sequence 310, App
C 122	14.4	80.0	60	15	US-10-272-437A-26	Sequence 26, Appl	195	14.4	80.0	3296	10	US-09-906-823-310	Sequence 310, App
C 123	14.4	80.0	174	15	US-10-029-386-19990	Sequence 19990, A	196	14.4	80.0	3296	10	US-09-907-613-310	Sequence 310, App
C 124	14.4	80.0	186	17	US-10-702-075-560	Sequence 560, App	197	14.4	80.0	3296	10	US-09-907-944-310	Sequence 310, App
C 125	14.4	80.0	192	17	US-10-437-963-4054	Sequence 4054, Ap	198	14.4	80.0	3296	10	US-09-904-853-310	Sequence 310, App
C 126	14.4	80.0	250	17	US-10-437-963-100379	Sequence 100379, A	199	14.4	80.0	3296	10	US-09-909-820-310	Sequence 310, App
C 127	14.4	80.0	257	18	US-10-425-115-44446	Sequence 44446, A	200	14.4	80.0	3296	10	US-09-904-820-310	Sequence 310, App
C 128	14.4	80.0	399	16	US-10-424-599-15793	Sequence 15793, A	201	14.4	80.0	3296	10	US-09-904-786-310	Sequence 310, App
C 129	14.4	80.0	424	15	US-10-425-115-16248	Sequence 16248, A	202	14.4	80.0	3296	10	US-09-906-646-310	Sequence 310, App
C 130	14.4	80.0	439	15	US-10-210-314-5	Sequence 5, Appl	203	14.4	80.0	3296	10	US-09-906-700-310	Sequence 310, App
C 131	14.4	80.0	458	18	US-10-425-115-40702	Sequence 40702, A	204	14.4	80.0	3296	10	US-09-903-785-310	Sequence 310, App
C 132	14.4	80.0	478	10	US-09-918-995-2414	Sequence 2414, A	205	14.4	80.0	3296	10	US-09-902-903-310	Sequence 310, App
C 133	14.4	80.0	493	10	US-09-918-995-31960	Sequence 31960, A	206	14.4	80.0	3296	10	US-09-903-749A-110	Sequence 310, App
C 134	14.4	80.0	502	17	US-10-021-323-1270	Sequence 1270, Ap	207	14.4	80.0	3296	10	US-09-904-119-310	Sequence 310, App
C 135	14.4	80.0	530	15	US-10-029-386-6259	Sequence 6259, Ap	208	14.4	80.0	3296	10	US-09-904-956-310	Sequence 310, App
C 136	14.4	80.0	532	17	US-10-767-701-3765	Sequence 3765, A	209	14.4	80.0	3296	10	US-09-902-736-310	Sequence 310, App
C 137	14.4	80.0	534	13	US-10-027-632-49968	Sequence 49968, A	210	14.4	80.0	3296	10	US-09-907-799-310	Sequence 310, App
C 138	14.4	80.0	553	15	US-10-027-632-49968	Sequence 49968, A	211	14.4	80.0	3296	10	US-09-903-943-310	Sequence 310, App
C 139	14.4	80.0	553	10	US-09-918-995-11478	Sequence 11478, A	212	14.4	80.0	3296	10	US-09-904-463-310	Sequence 310, App
C 140	14.4	80.0	569	13	US-10-027-632-81739	Sequence 81739, A	213	14.4	80.0	3296	10	US-09-907-922-310	Sequence 310, App
C 141	14.4	80.0	569	15	US-10-027-632-81739	Sequence 81739, A	214	14.4	80.0	3296	10	US-09-902-699-310	Sequence 310, App
C 142	14.4	80.0	591	17	US-10-479-334-2	Sequence 2, Appl	215	14.4	80.0	3296	10	US-09-903-520-310	Sequence 310, App
C 143	14.4	80.0	591	17	US-10-479-334-11	Sequence 11, Appl	216	14.4	80.0	3296	10	US-09-905-055-310	Sequence 310, App
C 144	14.4	80.0	596	16	US-10-424-599-99594	Sequence 99594, A	217	14.4	80.0	3296	10	US-09-909-064-310	Sequence 310, App
C 145	14.4	80.0	605	18	US-10-425-115-143653	Sequence 143653, A	218	14.4	80.0	3296	10	US-09-904-553-310	Sequence 310, App
C 146	14.4	80.0	650	16	US-10-424-599-9541	Sequence 9541, Ap	219	14.4	80.0	3296	10	US-09-905-381-310	Sequence 310, App
C 147	14.4	80.0	684	9	US-09-938-842A-963	Sequence 963, App	220	14.4	80.0	3296	10	US-09-904-485-310	Sequence 310, App
C 148	14.4	80.0	684	11	US-09-938-842A-963	Sequence 963, App	221	14.4	80.0	3296	10	US-09-905-348-310	Sequence 310, App
C 149	14.4	80.0	718	13	US-10-027-632-144355	Sequence 144355, A	222	14.4	80.0	3296	10	US-09-905-088-310	Sequence 310, App
C 150	14.4	80.0	718	15	US-10-027-632-144355	Sequence 144355, A	223	14.4	80.0	3296	10	US-09-907-575-310	Sequence 310, App
C 151	14.4	80.0	732	16	US-10-424-599-88952	Sequence 88952, A	224	14.4	80.0	3296	10	US-09-905-075-310	Sequence 310, App
C 152	14.4	80.0	748	17	US-10-437-963-83366	Sequence 83366, A	225	14.4	80.0	3296	10	US-09-902-759-310	Sequence 310, App
C 153	14.4	80.0	765	17	US-10-437-963-83366	Sequence 83366, A	226	14.4	80.0	3296	10	US-09-902-633-310	Sequence 310, App
C 154	14.4	80.0	802	16	US-10-424-599-5626	Sequence 5626, Ap	227	14.4	80.0	3296	10	US-09-902-713-310	Sequence 310, App
C 155	14.4	80.0	820	15	US-10-369-493-38770	Sequence 38770, A	228	14.4	80.0	3296	10	US-09-907-979-310	Sequence 310, App
C 156	14.4	80.0	893	17	US-10-437-963-18416	Sequence 18416, A	229	14.4	80.0	3296	10	US-09-902-615-310	Sequence 310, App
C 157	14.4	80.0	1126	15	US-10-369-493-32264	Sequence 32264, A	230	14.4	80.0	3296	10	US-09-903-925-310	Sequence 310, App
C 158	14.4	80.0	1159	11	US-09-985-153-45	Sequence 45, Appl	231	14.4	80.0	3296	10	US-09-906-760A-310	Sequence 310, App

232	14.4	80.0	3266	10	US-09-903-823-310	Sequence 310, App	305	14.4	80.0	3266	14	US-10-147-527-369	Sequence 369, App
233	14.4	80.0	3266	10	US-09-907-652-310	Sequence 310, App	306	14.4	80.0	3266	14	US-10-121-041-369	Sequence 369, App
234	14.4	80.0	3266	10	US-09-902-572A-310	Sequence 310, App	307	14.4	80.0	3266	14	US-10-121-043-369	Sequence 369, App
235	14.4	80.0	3266	10	US-09-902-979-310	Sequence 310, App	308	14.4	80.0	3266	14	US-10-121-047-369	Sequence 369, App
236	14.4	80.0	3266	10	US-09-905-125-310	Sequence 310, App	309	14.4	80.0	3266	14	US-10-123-015-369	Sequence 369, App
237	14.4	80.0	3266	10	US-09-906-815A-310	Sequence 310, App	310	14.4	80.0	3266	14	US-10-123-998-369	Sequence 369, App
238	14.4	80.0	3266	10	US-09-905-448-310	Sequence 310, App	311	14.4	80.0	3266	14	US-10-123-992-369	Sequence 369, App
239	14.4	80.0	3266	10	US-09-903-808-310	Sequence 310, App	312	14.4	80.0	3266	14	US-10-123-910-369	Sequence 369, App
240	14.4	80.0	3266	10	US-09-904-992-310	Sequence 310, App	313	14.4	80.0	3266	14	US-10-124-813-369	Sequence 369, App
241	14.4	80.0	3266	10	US-09-904-838-310	Sequence 310, App	314	14.4	80.0	3266	14	US-10-124-817-369	Sequence 369, App
242	14.4	80.0	3266	10	US-09-906-777-310	Sequence 310, App	315	14.4	80.0	3266	14	US-10-125-924-369	Sequence 369, App
243	14.4	80.0	3266	10	US-09-903-603A-310	Sequence 310, App	316	14.4	80.0	3266	14	US-10-125-924-369	Sequence 369, App
244	14.4	80.0	3266	10	US-09-904-533-310	Sequence 310, App	317	14.4	80.0	3266	14	US-10-125-924-369	Sequence 369, App
245	14.4	80.0	3266	10	US-09-904-766-310	Sequence 310, App	318	14.4	80.0	3266	14	US-10-140-660-369	Sequence 369, App
246	14.4	80.0	3266	10	US-09-904-922A-310	Sequence 310, App	319	14.4	80.0	3266	14	US-10-142-417-369	Sequence 369, App
247	14.4	80.0	3266	10	US-09-904-877A-310	Sequence 310, App	320	14.4	80.0	3266	14	US-10-147-519-369	Sequence 369, App
248	14.4	80.0	3266	10	US-09-903-562-310	Sequence 310, App	321	14.4	80.0	3266	14	US-10-157-762-369	Sequence 369, App
249	14.4	80.0	3266	10	US-09-906-618-310	Sequence 310, App	322	14.4	80.0	3266	14	US-10-152-355-369	Sequence 369, App
250	14.4	80.0	3266	10	US-09-907-728-310	Sequence 310, App	323	14.4	80.0	3266	14	US-10-125-966A-369	Sequence 369, App
251	14.4	80.0	3266	10	US-09-904-805-310	Sequence 310, App	324	14.4	80.0	3266	14	US-10-125-910A-369	Sequence 369, App
252	14.4	80.0	3266	10	US-09-904-938A-310	Sequence 310, App	325	14.4	80.0	3266	14	US-10-127-831A-369	Sequence 369, App
253	14.4	80.0	3266	10	US-09-906-722A-310	Sequence 310, App	326	14.4	80.0	3266	14	US-10-127-837A-369	Sequence 369, App
254	14.4	80.0	3266	11	US-09-908-578-310	Sequence 310, App	327	14.4	80.0	3266	14	US-10-127-882A-369	Sequence 369, App
255	14.4	80.0	3266	14	US-10-028-072-369	Sequence 369, App	328	14.4	80.0	3266	14	US-10-127-882A-369	Sequence 369, App
256	14.4	80.0	3266	14	US-10-140-808-369	Sequence 369, App	329	14.4	80.0	3266	14	US-10-127-843A-369	Sequence 369, App
257	14.4	80.0	3266	14	US-10-121-049-369	Sequence 369, App	330	14.4	80.0	3266	14	US-10-127-845A-369	Sequence 369, App
258	14.4	80.0	3266	14	US-10-123-904-369	Sequence 369, App</							

378	14.4	80.0	3296	14	US-10-128-664A-369	Sequence 369, App	451	14.4	80.0	3296	15	US-10-141-700-369	Sequence 369, App
379	14.4	80.0	3296	14	US-10-123-825A-369	Sequence 369, App	452	14.4	80.0	3296	15	US-10-141-705-369	Sequence 369, App
380	14.4	80.0	3296	14	US-10-230-417-369	Sequence 369, App	453	14.4	80.0	3296	15	US-10-141-753-369	Sequence 369, App
381	14.4	80.0	3296	14	US-10-131-815A-369	Sequence 369, App	454	14.4	80.0	3296	15	US-10-141-758-369	Sequence 369, App
382	14.4	80.0	3296	14	US-10-131-817A-369	Sequence 369, App	455	14.4	80.0	3296	15	US-10-142-118-369	Sequence 369, App
383	14.4	80.0	3296	14	US-10-131-821A-369	Sequence 369, App	456	14.4	80.0	3296	15	US-10-142-120-369	Sequence 369, App
384	14.4	80.0	3296	14	US-10-131-822A-369	Sequence 369, App	457	14.4	80.0	3296	15	US-10-142-122-369	Sequence 369, App
385	14.4	80.0	3296	14	US-10-131-828A-369	Sequence 369, App	458	14.4	80.0	3296	15	US-10-142-127-369	Sequence 369, App
386	14.4	80.0	3296	14	US-10-131-835A-369	Sequence 369, App	459	14.4	80.0	3296	15	US-10-142-160-369	Sequence 369, App
387	14.4	80.0	3296	14	US-10-131-864A-369	Sequence 369, App	460	14.4	80.0	3296	15	US-10-145-821-369	Sequence 369, App
388	14.4	80.0	3296	14	US-10-137-869A-369	Sequence 369, App	461	14.4	80.0	3296	15	US-10-152-531-369	Sequence 369, App
389	14.4	80.0	3296	14	US-10-147-523-369	Sequence 369, App	462	14.4	80.0	3296	15	US-10-127-840B-369	Sequence 369, App
390	14.4	80.0	3296	14	US-10-158-785-369	Sequence 369, App	463	14.4	80.0	3296	15	US-10-142-624-369	Sequence 369, App
391	14.4	80.0	3296	14	US-10-121-051-369	Sequence 369, App	464	14.4	80.0	3296	15	US-10-142-761-369	Sequence 369, App
392	14.4	80.0	3296	14	US-10-121-042-369	Sequence 369, App	465	14.4	80.0	3296	15	US-10-142-765-369	Sequence 369, App
393	14.4	80.0	3296	14	US-10-123-912-369	Sequence 369, App	466	14.4	80.0	3296	15	US-10-142-857-369	Sequence 369, App
394	14.4	80.0	3296	15	US-10-129-007-369	Sequence 369, App	467	14.4	80.0	3296	15	US-10-142-887-369	Sequence 369, App
395	14.4	80.0	3296	15	US-10-194-359-369	Sequence 369, App	468	14.4	80.0	3296	15	US-10-143-088-369	Sequence 369, App
396	14.4	80.0	3296	15	US-10-127-847A-369	Sequence 369, App	469	14.4	80.0	3296	15	US-10-143-034-369	Sequence 369, App
397	14.4	80.0	3296	15	US-10-175-590-369	Sequence 369, App	470	14.4	80.0	3296	15	US-10-143-116-369	Sequence 369, App
398	14.4	80.0	3296	15	US-10-137-866-369	Sequence 369, App	471	14.4	80.0	3296	15	US-10-143-117-369	Sequence 369, App
399	14.4	80.0	3296	15	US-10-146-726-369	Sequence 369, App	472	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
400	14.4	80.0	3296	15	US-10-146-727-369	Sequence 369, App	473	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
401	14.4	80.0	3296	15	US-10-146-768-369	Sequence 369, App	474	14.4	80.0	3296	15	US-10-145-015-369	Sequence 369, App
402	14.4	80.0	3296	15	US-10-152-380-369	Sequence 369, App	475	14.4	80.0	3296	15	US-10-145-090-369	Sequence 369, App
403	14.4	80.0	3296	15	US-10-153-934-369	Sequence 369, App	476	14.4	80.0	3296	15	US-10-145-091-369	Sequence 369, App
404	14.4	80.0	3296	15	US-10-140-807-369	Sequence 369, App							

[illegible]

670	14.4	80.0	3296	15	US-10-140-865-359	Sequence 369, App	743	14	77.8	875	16	US-10-424-599-29925	Sequence 29925, A
671	14.4	80.0	3296	15	US-10-141-701-369	Sequence 369, App	744	14	77.8	892	17	US-10-437-963-92960	Sequence 92960, A
672	14.4	80.0	3296	15	US-10-141-754-369	Sequence 369, App	745	14	77.8	1022	16	US-10-424-599-4182	Sequence 4182, A
673	14.4	80.0	3296	15	US-10-141-760-369	Sequence 369, App	746	14	77.8	1041	17	US-10-437-963-71730	Sequence 71730, A
674	14.4	80.0	3296	15	US-10-142-425-359	Sequence 369, App	747	14	77.8	1209	15	US-10-437-963-6634	Sequence 6634, A
675	14.4	80.0	3296	15	US-10-142-430-359	Sequence 369, App	748	14	77.8	1515	15	US-10-395-400-141	Sequence 141, App
676	14.4	80.0	3296	15	US-10-143-113-359	Sequence 369, App	749	14	77.8	1515	16	US-10-412-639B-535	Sequence 535, App
677	14.4	80.0	3296	15	US-10-146-730-359	Sequence 369, App	750	14	77.8	1653	16	US-10-282-122A-18952	Sequence 18952, A
678	14.4	80.0	3296	15	US-10-146-792-359	Sequence 369, App	751	14	77.8	2207	17	US-10-437-963-30562	Sequence 30562, A
679	14.4	80.0	3296	15	US-10-158-791-359	Sequence 369, App	752	14	77.8	2217	17	US-10-437-963-92950	Sequence 92950, A
680	14.4	80.0	3296	15	US-10-158-843-359	Sequence 369, App	753	14	77.8	2628	11	US-09-809-665A-19	Sequence 39, App1
681	14.4	80.0	3296	15	US-10-157-788-359	Sequence 369, App	754	14	77.8	2833	17	US-10-437-963-73163	Sequence 73163, A
682	14.4	80.0	3296	15	US-10-152-405-359	Sequence 369, App	755	14	77.8	3144	17	US-10-437-963-73167	Sequence 8080, A
683	14.4	80.0	3296	15	US-10-152-405-359	Sequence 369, App	756	14	77.8	4428	17	US-10-437-963-80480	Sequence 963, App
684	14.4	80.0	3296	15	US-10-147-528-359	Sequence 369, App	757	14	77.8	6855	15	US-10-017-161-963	Sequence 106, App
685	14.4	80.0	3296	15	US-10-148-923-310	Sequence 310, App	758	14	77.8	6678	18	US-10-739-930-106	Sequence 1015, App
686	14.4	80.0	3296	15	US-10-148-923-310	Sequence 310, App	759	14	77.8	27459	15	US-10-017-161-1015	Sequence 1744, App
687	14.4	80.0	3296	16	US-10-148-923-310	Sequence 310, App	760	14	77.8	53885	13	US-10-087-192-1142	Sequence 112, App
688	14.4	80.0	3296	16	US-10-148-923-310	Sequence 369, App	761	14	77.8	91278	17	US-10-322-281-112	Sequence 341, App
689	14.4	80.0	3296	16	US-10-140-927-359	Sequence 369, App	762	14	77.8	118931	13	US-10-087-192-1108	Sequence 1108, App
690	14.4	80.0	3296	16	US-10-140-927-359	Sequence 369, App	763	14	77.8	256190	17	US-10-322-281-330	Sequence 320, App
691	14.4	80.0	3296	16	US-10-147-493-359	Sequence 369, App	764	14	77.8	659158	9	US-09-771-208-20	Sequence 20, App1
692	14.4	80.0	3296	16	US-10-147-493-359	Sequence 369, App	765	14	77.8	2140405	13	US-10-027-632-76212	Sequence 76212, A
693	14.4	80.0	3296	16	US-10-147-493-359	Sequence 369, App	766	14	77.8	2140405	15	US-10-027-632-76212	Sequence 174763, A
694	14.4	80.0	3296	16	US-10-147-493-359	Sequence 369, App	767	14	77.8	240917	13	US-10-027-632-174763	Sequence 174763, A
695	14.4	80.0	3296	16	US-10-147-493-359	Sequence 369, App	768	14	77.8	240917	15	US-10-027-632-174763	Sequence 174763, A
696	14.4	80.0	3296	16	US-10-158-781-359	Sequence 369, App	769	14	77.8	240917	15	US-10-027-632-174763	Sequence 1643, App
697	14.4	80.0	3296	16	US-10-147-536-359	Sequence 369, App	770	13.8	76.7	66	8	US-08-781-986A-1543	Sequence 1543, App
698	14.4	80.0	3296	16	US-10-147-536-359	Sequence 369, App	771	13.8	76.7	66	6	US-10-329-624-1543	Sequence 1643, App
699	14.4	80.0	3296	17	US-10-152-372-359	Sequence 369, App	772	13.8	76.7	159	16	US-10-282-122A-1569	Sequence 1569, App
700	14.4	80.0	3296	17	US-10-152-372-359	Sequence 310, App	773	13.8	76.7	160	9	US-09-604-287A-44	Sequence 44, App1
701	14.4	80.0	3296	17	US-10-152-372-359	Sequence 310, App	774	13.8	76.7	160	9	US-09-604-287A-44	Sequence 44, App1
702	14.4	80.0	3296	17	US-10-152-372-359	Sequence 310, App	775	13.8	76.7	160	9	US-09-339-338-44	Sequence 44, App1
703	14.4	80.0	3296	18	US-10-145-626-359	Sequence 369, App	776	13.8	76.7	160	10	US-09-551-621-44	Sequence 44, App1
704	14.4	80.0	3296	18	US-10-145-626-359	Sequence 369, App	777	13.8	76.7	160	13	US-10-007-800-44	Sequence 44, App1
705	14.4	80.0	3296	18	US-10-145-819-359	Sequence 369, App	778	13.8	76.7	160	14	US-10-076-622-44	Sequence 44, App1
706	14.4	80.0	3296	18	US-10-145-819-359	Sequence 369, App	779	13.8	76.7	160	15	US-10-124-800-44	Sequence 44, App1
707	14.4	80.0	3296	18	US-10-147-518-359	Sequence 369, App	780	13.8	76.7	160	15	US-10-441-893-44	Sequence 44, App1
708	14.4	80.0	3296	18	US-10-147-518-359	Sequence 369, App	781	13.8	76.7	170	9	US-09-815-242-3038	Sequence 3028, App
709	14.4	80.0	3296	17	US-10-145-961-359	Sequence 369, App	782	13.8	76.7	170	6	US-10-282-122A-5592	Sequence 5592, App
710	14.4	80.0	3296	17	US-10-145-961-359	Sequence 78567, A	783	13.8	76.7	171	18	US-10-425-115-182683	Sequence 182683, App
711	14.4	80.0	3296	17	US-10-145-961-359	Sequence 12579, A	784	13.8	76.7	205	17	US-10-437-963-10642	Sequence 10642, App
712	14.4	80.0	3296	18	US-10-425-115-152026	Sequence 125026, A	785	13.8	76.7	223	9	US-09-983-965-1301	Sequence 1371, App
713	14.4	80.0	3296	17	US-10-425-115-152026	Sequence 33552, A	786	13.8	76.7	247	9	US-09-834-975-537	Sequence 537, App
714	14.4	80.0	4023	18	US-10-425-115-18862	Sequence 33862, A	787	13.8	76.7	250	18	US-10-425-115-145452	Sequence 145452, App
715	14.4	80.0	4426	16	US-10-094-886-123	Sequence 123, App	788	13.8	76.7	251	16	US-10-085-783A-14373	Sequence 14373, A
716	14.4	80.0	4488	8	US-08-781-986A-228	Sequence 228, App	789	13.8	76.7	251	16	US-10-085-783A-14373	Sequence 2537, App
717	14.4	80.0	4488	16	US-10-323-624-228	Sequence 125, App	790	13.8	76.7	255	16	US-10-282-122A-1597	Sequence 51, App1
718	14.4	80.0	4501	16	US-09-764-872-752	Sequence 752, App	791	13.8	76.7	267	14	US-10-046-433-10	Sequence 4858, App
719	14.4	80.0	12123	10	US-09-764-872-752	Sequence 843, App	792	13.8	76.7	267	14	US-10-046-433-10	Sequence 20047, A
720	14.4	80.0	35183	17	US-10-323-281-445	Sequence 843, App	793	13.8	76.7	288	9	US-09-983-965-4858	Sequence 136166, App
721	14.4	80.0	89213	17	US-10-323-281-445	Sequence 3, App1	794	13.8	76.7	305	15	US-10-029-368-20047	Sequence 2819, App
722	14.4	80.0	119556	14	US-10-270-336-3	Sequence 1, App1	795	13.8	76.7	317	16	US-10-282-122A-1819	Sequence 42451, App
723	14.4	80.0	162450	15	US-10-071-179-1	Sequence 1, App1	796	13.8	76.7	322	16	US-10-425-115-16293	Sequence 16293, A
724	14.4	80.0	162450	15	US-10-126-704-1	Sequence 3, App1	797	13.8	76.7	327	18	US-10-425-115-16293	Sequence 8114, A
725	14.4	80.0	397658	9	US-09-813-320-3	Sequence 4, App1	798	13.8	76.7	330	18	US-10-425-115-81314	Sequence 4180, App
726	14.4	80.0	513509	10	US-09-754-853A-4	Sequence 29057, A	799	13.8	76.7	336	9	US-09-815-242-4180	Sequence 3462, App
727	14.4	77.8	401	16	US-10-424-599-29057	Sequence 71009, A	800	13.8	76.7	331	16	US-10-282-122A-4362	Sequence 8160, App
728	14.4	77.8	462	17	US-10-424-599-29057	Sequence 359, App	801	13.8	76.7	334	9	US-09-815-242-810	Sequence 7951, App
729	14.4	77.8	503	10	US-09-770-961-359	Sequence 2505, App	802	13.8	76.7	334	9	US-10-282-122A-7931	Sequence 23339, App
730	14.4	77.8	532	17	US-10-653-047-603	Sequence 15364, A	803	13.8	76.7	334	16	US-10-424-599-116618	Sequence 11584, App
731	14.4	77.8	559	18	US-10-425-115-15364	Sequence 15364, A	804	13.8	76.7	334	16	US-10-424-599-116618	Sequence 156, App
732	14.4	77.8	577	18	US-10-369-493-31155	Sequence 203812, A	805	13.8	76.7	334	16	US-10-424-599-116618	Sequence 27667, A
733	14.4	77.8	627	15	US-10-369-493-31155	Sequence 203812, A	806	13.8	76.7	334	16	US-10-424-599-116618	Sequence 6715, App
734	14.4	77.8	637	15	US-10-369-493-31155	Sequence 4405, App	807	13.8	76.7	334	16	US-10-424-599-116618	Sequence 962, App
735	14.4	77.8	669	15	US-10-027-632-203813	Sequence 99215, A	808	13.8	76.7	411	9	US-09-834-975-155	Sequence 2597, App
736	14.4	77.8	669	13	US-10-027-632-203813	Sequence 203813, A	809	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
737	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	810	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
738	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	811	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
739	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	812	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
740	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	813	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
741	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	814	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A
742	14.4	77.8	669	15	US-10-027-632-203813	Sequence 203812, A	815	13.8	76.7	411	9	US-09-834-975-155	Sequence 28626, A

C 816	13.8	76.7	469	18	US-10-425-115-176565	Sequence 176565,
817	13.8	76.7	478	9	US-09-908-931B-47	Sequence 47, Appl
818	13.8	76.7	478	16	US-10-333-964-48	Sequence 48, Appl
819	13.8	76.7	480	16	US-10-260-238-5147	Sequence 5147, Ap
820	13.8	76.7	481	16	US-10-424-539-41021	Sequence 41021, A
821	13.8	76.7	486	16	US-10-260-238-1471	Sequence 1471, Ap
822	13.8	76.7	506	16	US-10-424-559-139527	Sequence 139527,
823	13.8	76.7	509	16	US-10-437-963-57351	Sequence 57351, A
824	13.8	76.7	512	15	US-10-029-386-6316	Sequence 6316, Ap
825	13.8	76.7	519	15	US-10-027-632-92835	Sequence 92835, A
826	13.8	76.7	519	15	US-10-027-632-92835	Sequence 92835, A
C 827	13.8	76.7	520	18	US-10-425-115-14778	Sequence 14778, A
C 828	13.8	76.7	529	17	US-10-767-701-2908	Sequence 2908, Ap
C 829	13.8	76.7	531	9	US-09-770-152-822	Sequence 822, App
830	13.8	76.7	545	16	US-10-424-559-114152	Sequence 114152,
831	13.8	76.7	546	15	US-10-425-114-31358	Sequence 31358, A
832	13.8	76.7	558	16	US-10-161-051-46	Sequence 46, Appl
833	13.8	76.7	580	15	US-10-161-051-46	Sequence 89488, A
C 834	13.8	76.7	586	17	US-10-425-115-152937	Sequence 152937,
835	13.8	76.7	588	18	US-10-027-633-6730	Sequence 6730, A
836	13.8	76.7	589	13	US-10-027-633-6730	Sequence 6730, A
837	13.8	76.7	589	13	US-10-027-633-6730	Sequence 6730, A
838	13.8	76.7	589	13	US-10-027-633-6730	Sequence 6730, A
839	13.8	76.7	589	15	US-10-027-633-6730	Sequence 6730, A
840	13.8	76.7	589	15	US-10-027-633-6730	Sequence 6730, A
841	13.8	76.7	589	15	US-10-027-633-6730	Sequence 6730, A
842	13.8	76.7	590	17	US-10-021-323-311242	Sequence 311242,
843	13.8	76.7	590	17	US-10-021-323-311242	Sequence 311242,
844	13.8	76.7	605	18	US-10-425-115-6950	Sequence 6950, Ap
845	13.8	76.7	619	9	US-09-796-692-6548	Sequence 6548, Ap
846	13.8	76.7	619	14	US-10-040-863-6548	Sequence 6548, Ap
847	13.8	76.7	619	16	US-10-057-478B-6548	Sequence 6548, Ap
848	13.8	76.7	619	16	US-10-154-884B-6548	Sequence 6548, Ap
849	13.8	76.7	619	17	US-10-764-324-6548	Sequence 6548, Ap
850	13.8	76.7	634	17	US-10-767-701-1433	Sequence 1433, Ap
851	13.8	76.7	634	17	US-10-198-846-9002	Sequence 9002, Ap
852	13.8	76.7	634	13	US-10-027-633-209173	Sequence 209173,
C 853	13.8	76.7	637	15	US-10-027-633-209173	Sequence 209173,
C 854	13.8	76.7	640	16	US-10-425-114-279	Sequence 279, App
C 855	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 856	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 857	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 858	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 859	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 860	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 861	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 862	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 863	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 864	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 865	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 866	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 867	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 868	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 869	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 870	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 871	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 872	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 873	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 874	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 875	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 876	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 877	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 878	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 879	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 880	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 881	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 882	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 883	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 884	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 885	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 886	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 887	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A
C 888	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16159, A

```
962 13.8 76.7 1599 15 US-10-128-714-2442 Sequence 2442, Ap
963 13.8 76.7 1599 15 US-10-128-714-2442 Sequence 2442, Ap
964 13.8 76.7 1614 16 US-10-425-114-14594 Sequence 14594, A
965 13.8 76.7 1639 18 US-10-425-115-56500 Sequence 56500, A
966 13.8 76.7 1643 16 US-10-425-114-26886 Sequence 26886, A
967 13.8 76.7 1643 18 US-10-425-115-75285 Sequence 75285, A
968 13.8 76.7 1670 18 US-10-425-115-178309 Sequence 178309,
969 13.8 76.7 1680 9 US-09-071-035-85 Sequence 85, Appl
970 13.8 76.7 1680 14 US-10-206-576-85 Sequence 85, Appl
971 13.8 76.7 1705 16 US-10-425-114-6395 Sequence 6395, Ap
972 13.8 76.7 1716 15 US-10-369-493-34585 Sequence 34585, A
973 13.8 76.7 1726 17 US-10-767-701-13885 Sequence 13885, A
974 13.8 76.7 1739 15 US-10-128-714-1442 Sequence 1442, Ap
975 13.8 76.7 1739 15 US-10-128-714-6442 Sequence 6442, Ap
976 13.8 76.7 1749 9 US-09-938-842A-3369 Sequence 3369, Ap
977 13.8 76.7 1749 11 US-09-938-842A-3369 Sequence 3369, Ap
978 13.8 76.7 1760 18 US-10-425-115-90240 Sequence 90240, A
979 13.8 76.7 1762 18 US-10-425-115-178311 Sequence 178311,
980 13.8 76.7 1819 14 US-10-219-220-243 Sequence 243, App
981 13.8 76.7 1829 9 US-09-764-864-190 Sequence 190, App
982 13.8 76.7 1829 18 US-10-739-930-2921 Sequence 2921, Ap
983 13.8 76.7 1833 16 US-10-425-114-28491 Sequence 28491, A
984 13.8 76.7 1834 16 US-10-424-599-90755 Sequence 90755, A
985 13.8 76.7 1846 18 US-10-739-930-4328 Sequence 4328, A
986 13.8 76.7 1854 15 US-10-037-270-793 Sequence 793, App
987 13.8 76.7 1854 15 US-10-117-722-793 Sequence 793, App
988 13.8 76.7 1874 10 US-09-822-846-188 Sequence 188, App
989 13.8 76.7 1883 16 US-10-374-780A-1759 Sequence 1759, Ap
990 13.8 76.7 1922 18 US-10-425-115-178933 Sequence 178933,
991 13.8 76.7 1927 18 US-10-425-115-178937 Sequence 178937,
992 13.8 76.7 1949 17 US-10-115-635-55 Sequence 55, Appl
993 13.8 76.7 1965 16 US-10-282-122A-7372 Sequence 7372, Ap
994 13.8 76.7 1995 18 US-10-425-115-56269 Sequence 56269, A
995 13.8 76.7 2000 9 US-09-887-576-147 Sequence 147, App
996 13.8 76.7 2000 9 US-09-938-842A-4847 Sequence 4847, Ap
997 13.8 76.7 2000 11 US-09-938-842A-4847 Sequence 4847, Ap
998 13.8 76.7 2027 9 US-09-771-161A-29 Sequence 29, Appl
999 13.8 76.7 2055 14 US-10-004-633-3 Sequence 3, Appl
1000 13.8 76.7 2055 15 US-10-252-131-3 Sequence 3, Appl
```

ALIGNMENTS

```
RESULT 1
US-10-050-189a-6
; Sequence 6, Application US/10050189a
; Publication No. US2002016856A1
; GENERAL INFORMATION:
; APPLICANT: Rubin, Berish
; TITLE OF INVENTION: Detection of Mutations in a Gene Encoding IKB Kinase-Complex-Asso
; FILE REFERENCE: Rubin 201
; CURRENT APPLICATION NUMBER: US/10/050,189A
; CURRENT FILING DATE: 2002-01-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-050-189a-6
```

```
Query Match 100.0%; Score 18; DB 13; Length 18;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GAGAACACACAGATTCG 18
Db 1 GAGAACACACAGATTCG 18
```

```
RESULT 2
US-10-027-632-111967/c
; Sequence 111967, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111967
; LENGTH: 2465
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-111967
```

```
Query Match 88.9%; Score 16; DB 13; Length 2465;
Best Local Similarity 100.0%; Pred. No. 58+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GAGAACACACAGATTC 16
Db 1560 GAGAACACACAGATTC 1545
```

```
RESULT 3
US-10-027-632-111967/c
; Sequence 111967, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111967
; LENGTH: 2465
; TYPE: DNA
; ORGANISM: Human
```

US-10-027-632-111967

Query Match 88.9%; Score 16; DB 15; Length 2465;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTTC 16
Db 1560 GAGAACACACAGATTTC 1545

RESULT 4
US-10-041-856-1
; Sequence 1, Application US/10041856
; Publication No. US20020169299A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENHAUPT, SUSAN
; APPLICANT: GUSELLA, JAMES P.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; TITLE OF INVENTION: DYSAUTONOMIA
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/10/041,856
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: 60/260,080
; PRIOR FILING DATE: 2001-01-06
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 66479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-041-856-1

Query Match 88.9%; Score 16; DB 13; Length 66479;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTTC 16
Db 33979 GAGAACACACAGATTTC 33994

RESULT 5
US-09-833-381-908
; Sequence 908, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 908
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-908

Query Match 85.6%; Score 15.4; DB 9; Length 409;
Best Local Similarity 94.1%; Pred. No. 8.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTTCG 18
Db 170 AGAACACACAGATTTCG 186

RESULT 6
US-09-878-178-1155
; Sequence 1155, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 21021.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-1155

Query Match 85.6%; Score 15.4; DB 9; Length 482;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTTCG 18
Db 48 AGAACACACAGATTTCG 64

RESULT 7
US-10-046-935-1155
; Sequence 1155, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 21021.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-1155

Query Match 85.6%; Score 15.4; DB 13; Length 482;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTTCG 18
Db 48 AGAACACACAGATTTCG 64

RESULT 8
US-10-146-502-1155
; Sequence 1155, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun

```

; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-1155

Query Match      85.6%; Score 15.4; DB 14; Length 462;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
         |||
Db      48 AGCACACACAGATTCCG 64

RESULT 9
US-10-282-122A-20772
; Sequence 20772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/151,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20772
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-282-122A-20772
```

```

Query Match      85.6%; Score 15.4; DB 16; Length 513;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
         |||
Db      447 AGAACACACAGATTCCG 463

RESULT 10
US-10-027-632-277903/c
; Sequence 277903, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277903
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-277903

Query Match      85.6%; Score 15.4; DB 13; Length 577;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
         |||
Db      31 AGCACACACAGATTCCG 15

RESULT 11
US-10-027-632-277903/c
; Sequence 277903, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 277903
;; LENGTH: 577
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-277903

Query Match 85.6%; Score 15.4; DB 15; Length 577;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 31 AGCACACACAGATTGCG 15

RESULT 12
US-10-027-632-277719
; Sequence 277719, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277719
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-277719

Query Match 85.6%; Score 15.4; DB 13; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 553 AGCACACACAGATTGCG 569

RESULT 13
US-10-027-632-277719
; Sequence 277719, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 277719
;; LENGTH: 582
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-277719

Query Match 85.6%; Score 15.4; DB 15; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 553 AGCACACACAGATTGCG 569

RESULT 14
US-10-767-701-3256/C
; Sequence 3256, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5355)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 3256
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CTUS55082_1
US-10-767-701-3256

Query Match 85.6%; Score 15.4; DB 17; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 406 AGAACACACATATTGCG 390

RESULT 15
US-10-437-963-32919/C
; Sequence 32919, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.

```

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 32919
; LENGTH: 775
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37082C.1
US-10-437-963-32919

Query Match      85.6%; Score 15.4; DB 17; Length 775;
Best Local Similarity 94.1%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      598 AGAAGTACACAGATTCCG 582

RESULT 16
US-09-815-242-6118/c
; Sequence 6118, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6118
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1470)
US-09-815-242-6118

Query Match      85.6%; Score 15.4; DB 9; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
```

```

Db      757 AGAACACACAGATTCCG 741

RESULT 17
US-10-369-493-24585/c
; Sequence 24585, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Cao, Yongwei
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24585
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24585

Query Match      85.6%; Score 15.4; DB 15; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      757 AGAACACACAGATTCCG 741

RESULT 18
US-10-369-493-32483
; Sequence 32483, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32483
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Ralstonia metallidurans
US-10-369-493-32483

Query Match      85.6%; Score 15.4; DB 15; Length 1518;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      1043 AGTACACACAGATTCCG 1059

RESULT 19
US-10-369-493-45765
; Sequence 45765, Application US/10369493
```

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10 (52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45765
; LENGTH: 1950
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-45765

Query Match      85.6%; Score 15.4; DB 15; Length 1950;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACAACAAGATTGCG 18
Db      1795 AGAAGACAAGATTGCG 1811

RESULT 20
US-10-169-257-8/c
; Sequence 8, Application US/10169257
; Publication No. US20030207410A1
; GENERAL INFORMATION:
; APPLICANT: Tadashi FUJII
; APPLICANT: Yasuhide ARITOKU
; APPLICANT: Manabu MUKAIHARA
; APPLICANT: Takao NARITA
; APPLICANT: Hitosi AGEMATU
; APPLICANT: Kunio ISSHITI
; TITLE OF INVENTION: Biological process for the production of L-pipecolic acid
; FILE REFERENCE: 2002-0845A/WMC/00202
; CURRENT APPLICATION NUMBER: US/10/169,257
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: JP11/373389
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 20
; SEQ ID NO 8
; LENGTH: 2186
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
US-10-169-257-8

Query Match      85.6%; Score 15.4; DB 15; Length 2186;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACAACAAGATTGCG 18
Db      1281 AGAACAACAAGATTGCG 1265

RESULT 21
US-09-308-207-7
; Sequence 7, Application US/09308207
; Publication No. US2003022323A1
; GENERAL INFORMATION:
; APPLICANT: MARIA DIAZ-TORRES ET AL.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; PRODUCTION OF 1,3 PROPYLADIPOL
; NUMBER OF SEQUENCES: 68
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/308,207
; FILING DATE: 13-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: GUT2
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-308-207-7

Query Match      85.6%; Score 15.4; DB 10; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACAACAAGATTGCG 18
Db      2154 AGAACAACAAGATTGCG 2170

RESULT 22
US-10-093-463-135
; Sequence 135, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangoli, Esna
; APPLICANT: Verneet, Corine
; APPLICANT: Pena, Carol
```

APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytak, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Paturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taulier, Raymond J. Jr.
TITLE OF INVENTION: No. US2003020803A1el Antibodies that Bind to Antigenic Polypeptide
FILE REFERENCE: Encoding The Antigens, and Methods of Use.
CURRENT FILING DATE: 2002-06-24
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 135
LENGTH: 5101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (77)..(3308)
US-10-093-463-135
Query Match 85.6%; Score 15.4; DB 15; Length 5101;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

US-10-367-094-100
Sequence 100, Application US/10367094
Publication No. US20040170982A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
FILE REFERENCE: 529452001500
CURRENT APPLICATION NUMBER: US/10/367,094
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 203
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 100
LENGTH: 5719
TYPE: DNA
ORGANISM: Homo sapiens
US-10-367-094-100
Query Match 85.6%; Score 15.4; DB 17; Length 5719;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACAACAAGATTGCG 18
DB 368 AGCACAACAAGATTGCG 384
RESULT 24
US-09-070-927A-288
Sequence 288, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 7043 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 288;
US-09-070-927a-288

Query Match 85.6%; Score 15.4; DB 9; Length 7043;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTCCG 18
DB 1018 AGAACAACAAGATTCCG 1034

RESULT 25
US-10-367-094-99
Sequence 99, Application US/10367094
Publication No. US20040170982A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc Malandro
TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
FILE REFERENCE: 529452001500
CURRENT APPLICATION NUMBER: US/10/367,094
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 203
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 99
LENGTH: 142976
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(142976)
OTHER INFORMATION: n = A,T,C or G
US-10-367-094-99

Query Match 85.6%; Score 15.4; DB 17; Length 142976;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTCCG 18
DB 93623 AGAACAACAAGATTCCG 93639

RESULT 26
US-10-437-963-86110
Sequence 86110, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 86110
LENGTH: 378
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_85182C.1
US-10-437-963-86110

Query Match 83.3%; Score 15; DB 17; Length 378;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACAAGATT 15
DB 362 GAGAACACAAGATT 376

RESULT 27
US-10-137-036-4/C
Sequence 4, Application US/10137036
Publication No. US20030101478A1
GENERAL INFORMATION:
APPLICANT: Pereira, Ranjan
APPLICANT: Rice, Stephen
APPLICANT: Eagleton, Clare
APPLICANT: Lashaw, Annette
APPLICANT: Wood, Marion
APPLICANT: Visser, Elizabeth
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Modification of Gene Expression
FILE REFERENCE: 11000.1036c4
CURRENT APPLICATION NUMBER: US/10/137,036
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: PCT/NZ00/00018
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599
NUMBER OF SEQ ID NOS: 143
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 485
TYPE: DNA
ORGANISM: Pinus radiata
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(431)
FEATURE:
NAME/KEY: TATA_signal
LOCATION: (350)...(356)
FEATURE:
NAME/KEY: CAAT_signal
LOCATION: (326)...(333)
US-10-137-036-4

Query Match 83.3%; Score 15; DB 15; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACAACAAGATTCCG 18
DB 409 AACAACAAGATTCCG 395

RESULT 28
US-09-770-149-77/C
Sequence 77, Application US/09770149
Patent No. US20020059663A1
GENERAL INFORMATION:
APPLICANT: Goriach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thaliana
FILE REFERENCE: 2024 (PARA-013PRV)
CURRENT APPLICATION NUMBER: US/09/770,149
CURRENT FILING DATE: 2001-01-26
PRIORITY FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 77
LENGTH: 737
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(737)
OTHER INFORMATION: n = A,T,C or G
US-09-770-149-77

Query Match 83.3%; Score 15; DB 9; Length 737;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATT 15
DB 667 GAGACACACAGATT 653

RESULT 29
US-10-767-701-11437/c
Sequence 11437, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 11437
LENGTH: 1169
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS405_1
US-10-767-701-11437

Query Match 83.3%; Score 15; DB 17; Length 1169;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATT 16
DB 309 AGAACACACAGATT 295

RESULT 30
US-10-032-585-6133

Sequence 6133, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Busssey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6133
LENGTH: 1455
TYPE: DNA
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: misc_feature
LOCATION: (237)..(237)
OTHER INFORMATION: n=g, a, t or c
NAME/KEY: misc_feature
LOCATION: (348)..(348)
OTHER INFORMATION: n=g, a, t or c
US-10-032-585-6133

Query Match 83.3%; Score 15; DB 15; Length 1455;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATT 16
DB 801 AGAACACACAGATT 815

RESULT 31
US-10-425-115-87011
Sequence 87011, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53522)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 87011
LENGTH: 209
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_17935C.1
US-10-425-115-87011

Query Match 82.2%; Score 14.8; DB 18; Length 209;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGACACACAGATT 18
DB 157 GAGACACACAGATT 174

RESULT 32
US-09-867-550-1599/c
Sequence 1599, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.

```

; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1599
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-867-550-1599

Query Match      82.2%; Score 14.8; DB 9; Length 352;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGAACACAAAGATTGCG 18
DB      20 GAAACACAAAGATTGCG 3

RESULT 33
US-10-425-115-91075/C
; Sequence 91075, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 91075
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_183061C.1
; US-10-425-115-91075

Query Match      82.2%; Score 14.8; DB 18; Length 416;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGAACACAAAGATTGCG 18
DB      234 GAGACACAAAGATTGCG 217

RESULT 34
US-10-282-122A-33603
; Sequence 33603, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
```

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33603
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
; US-10-282-122A-33603

Query Match      82.2%; Score 14.8; DB 16; Length 438;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGAACACAAAGATTGCG 18
DB      85 GAGAACACAAAGATTGCG 102

RESULT 35
US-10-425-114-3291
; Sequence 3291, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3291
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258230_FLI
; US-10-425-114-3291

Query Match      82.2%; Score 14.8; DB 16; Length 452;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
```

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GAGAACACAGATTGCG 18
Db 187 GAGAACACAGATTGCG 204

RESULT 36
US-10-767-701-26974
; Sequence 26974, Application US/10767701
; Publication No. US2004017684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 26974
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 6672780
US-10-767-701-26974

Query Match 82.2%; Score 14.8; DB 17; Length 462;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 171 GAGACCAACAGTTGCG 188

RESULT 37
US-09-918-995-1845
; Sequence 1845, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1845
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1845

Query Match 82.2%; Score 14.8; DB 10; Length 465;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 329 GAGAAAACACATTGCG 346

RESULT 38

US-10-260-238-4785/c
; Sequence 4785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Krebs, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4785
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Musa acuminata
US-10-260-238-4785

Query Match 82.2%; Score 14.8; DB 16; Length 467;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 104 GAGAACACAGCTTACG 87

RESULT 39
US-10-425-115-81676/c
; Sequence 81676, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81676
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: M714577_174503C.1
US-10-425-115-81676

Query Match 82.2%; Score 14.8; DB 18; Length 470;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 297 GAAACACAGATTGCG 280

```

RESULT 40
US-10-027-632-68850/c
; Sequence 68850, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68850
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68850

```

```

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GAGACAAACAAGATTGCGC 18
DB 248 GAGATCAACAAGATGCGC 231

```

```

RESULT 41
US-10-027-632-68851/c
; Sequence 68851, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68851
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human

```

```

US-10-027-632-68851
Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GAGACAAACAAGATTGCGC 18
DB 248 GAGATCAACAAGATGCGC 231

```

```

RESULT 42
US-10-027-632-68850/c
; Sequence 68850, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68850
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68850

```

```

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GAGACAAACAAGATTGCGC 18
DB 248 GAGATCAACAAGATGCGC 231

```

```

RESULT 43
US-10-027-632-68851/c
; Sequence 68851, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

```

PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 68851
LENGTH: 473
TYPE: DNA
ORGANISM: Human
US-10-027-632-68851

Query Match 82.2%; Score 14.8; DB 15; Length 473;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 248 GAGATCAACAAGATCGCG 231

RESULT 44
US-10-424-599-16732
Sequence 16732, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 16732
LENGTH: 480
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_115114C.1
US-10-424-599-16732

Query Match 82.2%; Score 14.8; DB 16; Length 480;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 148 GAGAACCAACAAGCTTCC 165

RESULT 45
US-10-424-599-85425/C
Sequence 85425, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 85425
LENGTH: 498
TYPE: DNA
ORGANISM: Glycine max

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_48153C.1
US-10-424-599-85425

Query Match 82.2%; Score 14.8; DB 16; Length 498;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 470 GAGTACCAACAAGATCGCG 453

RESULT 46
US-10-027-632-217259/C
Sequence 217259, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108627.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 217259
LENGTH: 585
TYPE: DNA
ORGANISM: Human
US-10-027-632-217259

Query Match 82.2%; Score 14.8; DB 13; Length 585;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 409 GAGAACCAACAAGATTCCG 392

RESULT 47
US-10-027-632-217260/C
Sequence 217260, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108627.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 217260
LENGTH: 585
TYPE: DNA
ORGANISM: Human
US-10-027-632-217260

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 585;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTCCG 18
|||||
Db 409 GAGAACACACAGATTCCC 392

RESULT 48
US-10-027-632-217259/c
Sequence 217259, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 217259
LENGTH: 585
TYPE: DNA
ORGANISM: Human
US-10-027-632-217259

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 585;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTCCG 18
|||||
Db 409 GAGAACACACAGATTCCC 392

RESULT 49
US-10-027-632-217260/c
Sequence 217260, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 217260
LENGTH: 585
TYPE: DNA
ORGANISM: Human
US-10-027-632-217260

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 585;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTCCG 18
|||||
Db 409 GAGAACACACAGATTCCC 392

RESULT 50
US-10-425-114-5996
Sequence 5996, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO: 5996
LENGTH: 666
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700550605_FLI
US-10-425-114-5996

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 16; Length 666;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTCCG 18
|||||
Db 401 GAGAACACACAGATTCCG 418

Search completed: December 3, 2004, 07:42:25
Job time : 1601.42 secs

This Page Blank (uspio)

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 01:45:05 | Search time 59.6842 seconds
(without alignments)
214.365 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18

Sequence: 1 gagacaacacagatcgc 18

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backflten.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	ID	Description
1	15.4	85.6	534	US-09-134-000C-2523	Sequence 2523, App
2	15.4	85.6	3178	US-08-968-563-7	Sequence 7, App1
3	15.4	85.6	3178	US-08-969-683A-7	Sequence 7, App1
4	15.4	85.6	3178	US-09-297-928-3	Sequence 3, App1
5	15.4	85.6	485	US-09-276-599-4	Sequence 4, App1
6	15.4	85.6	485	US-09-598-401C-4	Sequence 4, App1
7	15.4	85.6	1359	US-09-248-796A-6542	Sequence 6542, App
8	14.8	82.2	676	US-09-495-050A-15	Sequence 15, App1
9	14.8	82.2	1098	US-09-320-757-10453	Sequence 10453, A
10	14.8	82.2	2727	US-09-328-352-3042	Sequence 3042, App
11	14.8	82.2	44377	US-08-804-158-1	Sequence 7, App1
12	14.8	82.2	44377	US-08-804-158-1	Sequence 7, App1
13	14.8	82.2	44377	US-08-804-158-1	Sequence 7, App1
14	14.8	82.2	44377	US-08-804-158-1	Sequence 7, App1
15	14.8	82.2	44377	US-08-804-158-1	Sequence 7, App1
16	14.4	80.0	591	US-09-252-991A-2843	Sequence 2843, App
17	14.4	80.0	684	US-09-252-991A-2843	Sequence 2843, App
18	14.4	80.0	771	US-09-270-757-13143	Sequence 13143, A
19	14.4	80.0	783	US-09-577-934A-1	Sequence 1, App1
20	14.4	80.0	867	US-09-252-991A-3123	Sequence 3123, App
21	14.4	80.0	1462	US-08-956-171E-771	Sequence 771, App
22	14.4	80.0	1462	US-08-956-171E-771	Sequence 771, App
23	14.4	80.0	1823	US-09-620-312D-516	Sequence 516, App
24	14.4	80.0	2853	US-09-556-601-25	Sequence 25, App1
25	14.4	80.0	3296	US-09-907-794A-310	Sequence 310, App
26	14.4	80.0	3296	US-09-907-794A-310	Sequence 310, App
27	14.4	80.0	3296	US-09-907-794A-310	Sequence 310, App

28	14.4	80.0	3296	US-09-906-700-310	Sequence 310, App
29	14.4	80.0	3296	US-10-140-002-369	Sequence 369, App
30	14.4	80.0	3296	US-09-903-603A-310	Sequence 310, App
31	14.4	80.0	4488	US-08-956-171E-228	Sequence 228, App
32	14.4	80.0	4488	US-08-956-171E-228	Sequence 228, App
33	14.4	80.0	162450	US-09-345-882-1	Sequence 1, App1
34	14.4	80.0	4403765	US-09-103-840A-2	Sequence 2, App1
35	14.4	80.0	4411529	US-09-252-991A-15780	Sequence 15780, A
36	14.4	77.8	723	US-09-248-796A-652	Sequence 652, App
37	14.4	77.8	723	US-09-248-796A-652	Sequence 652, App
38	14.4	77.8	835	US-09-252-991A-15910	Sequence 15910, A
39	14.4	77.8	835	US-09-252-991A-15910	Sequence 15910, A
40	14.4	77.8	1093	US-09-252-991A-15881	Sequence 15881, A
41	14.4	77.8	1317	US-09-252-991A-15881	Sequence 15881, A
42	14.4	77.8	2058	US-09-252-991A-15811	Sequence 15811, A
43	14.4	77.8	2628	US-09-609-665A-39	Sequence 39, App1
44	13.8	76.7	66	US-08-956-171E-1643	Sequence 1643, App
45	13.8	76.7	66	US-08-956-171E-1643	Sequence 1643, App
46	13.8	76.7	160	US-09-222-575-44	Sequence 44, App1
47	13.8	76.7	160	US-09-222-575-44	Sequence 44, App1
48	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
49	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
50	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
51	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
52	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
53	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
54	13.8	76.7	160	US-09-329-338-44	Sequence 44, App1
55	13.8	76.7	223	US-09-513-999C-16358	Sequence 16358, A
56	13.8	76.7	223	US-09-513-999C-16358	Sequence 16358, A
57	13.8	76.7	289	US-09-313-294A-2757	Sequence 2757, App
58	13.8	76.7	301	US-09-313-294A-2757	Sequence 2757, App
59	13.8	76.7	333	US-09-107-532A-1205	Sequence 1205, App
60	13.8	76.7	333	US-09-107-532A-1205	Sequence 1205, App
61	13.8	76.7	432	US-09-134-001C-146	Sequence 146, App
62	13.8	76.7	543	US-09-252-991A-2228	Sequence 2228, App
63	13.8	76.7	543	US-09-252-991A-2228	Sequence 2228, App
64	13.8	76.7	546	US-09-270-757-14473	Sequence 14473, A
65	13.8	76.7	723	US-09-601-198-161	Sequence 161, App
66	13.8	76.7	733	US-09-270-757-1287	Sequence 1287, App
67	13.8	76.7	859	US-09-270-757-1287	Sequence 1287, App
68	13.8	76.7	859	US-09-270-757-1287	Sequence 1287, App
69	13.8	76.7	933	US-09-252-991A-2302	Sequence 2302, App
70	13.8	76.7	933	US-09-252-991A-2302	Sequence 2302, App
71	13.8	76.7	978	US-09-270-757-23112	Sequence 23112, A
72	13.8	76.7	999	US-09-107-532A-1888	Sequence 1888, App
73	13.8	76.7	1002	US-09-134-000C-2256	Sequence 2256, App
74	13.8	76.7	1041	US-09-071-035-357	Sequence 357, App
75	13.8	76.7	1144	US-09-252-991A-2610	Sequence 2610, App
76	13.8	76.7	1144	US-09-252-991A-2610	Sequence 2610, App
77	13.8	76.7	1144	US-09-190-982-1	Sequence 1, App1
78	13.8	76.7	1271	US-09-408-257-1	Sequence 1, App1
79	13.8	76.7	1302	US-09-270-757-11058	Sequence 11058, A
80	13.8	76.7	1359	US-09-252-991A-12353	Sequence 12353, A
81	13.8	76.7	1392	US-09-134-001C-1136	Sequence 1136, App
82	13.8	76.7	1419	US-09-134-001C-1136	Sequence 1136, App
83	13.8	76.7	1589	US-09-248-796A-4402	Sequence 4402, App
84	13.8	76.7	1589	US-09-248-796A-4402	Sequence 4402, App
85	13.8	76.7	1611	US-09-071-035-87	Sequence 87, App1
86	13.8	76.7	1611	US-09-071-035-87	Sequence 87, App1
87	13.8	76.7	1677	US-09-270-757-12423	Sequence 12423, A
88	13.8	76.7	1677	US-09-270-757-12423	Sequence 12423, A
89	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
90	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
91	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
92	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
93	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
94	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
95	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
96	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
97	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
98	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
99	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A
100	13.8	76.7	1680	US-09-252-991A-12304	Sequence 12304, A

C 101	13.8	76.7	2055	4	US-09-130-367-3	Sequence 3, Appl	C 174	13.4	74.4	1806	4	US-09-134-000C-2378	Sequence 2378, Ap
C 102	13.8	76.7	2055	4	US-09-032-337-3	Sequence 3, Appl	C 175	13.4	74.4	1809	4	US-09-248-796A-6111	Sequence 6111, Ap
C 103	13.8	76.7	2055	4	US-09-464-231-3	Sequence 3	C 176	13.4	74.4	1815	4	US-09-248-796A-6111	Sequence 2421, Ap
C 104	13.8	76.7	2051	4	US-08-748-796A-12041	Sequence 12041, A	C 177	13.4	74.4	1854	4	US-09-023-655-1189	Sequence 1189, Ap
C 105	13.8	76.7	2082	1	US-08-178-477B-42	Sequence 42, Appl	C 178	13.4	74.4	1884	4	US-09-023-655-1189	Sequence 1189, Ap
C 106	13.8	76.7	3294	4	US-09-107-532A-99	Sequence 99, Appl	C 179	13.4	74.4	2025	3	US-08-932-0128-23	Sequence 23, Appl
C 107	13.8	76.7	3280	4	US-09-710-279-4337	Sequence 4337, Ap	C 180	13.4	74.4	2052	3	US-08-630-916A-45	Sequence 45, Appl
C 108	13.8	76.7	3386	1	US-09-710-279-3535	Sequence 3535, Ap	C 181	13.4	74.4	2052	3	US-09-657-481A-3	Sequence 3, Appl
C 109	13.8	76.7	3429	1	US-08-097-997A-10	Sequence 10, Appl	C 182	13.4	74.4	2106	4	US-09-710-279-109	Sequence 709, Appl
C 110	13.8	76.7	3429	3	US-08-665-574C-10	Sequence 10, Appl	C 183	13.4	74.4	2106	4	US-09-710-279-109	Sequence 709, Appl
C 111	13.8	76.7	3429	3	US-08-946-994-10	Sequence 10, Appl	C 184	13.4	74.4	2127	4	US-09-252-991A-15511	Sequence 15511, Ap
C 112	13.8	76.7	3501	4	US-10-140-002-37	Sequence 37, Appl	C 185	13.4	74.4	2409	1	US-07-911-531-18	Sequence 18, Appl
C 113	13.8	76.7	3588	4	US-09-963-137-201	Sequence 201, App	C 186	13.4	74.4	2409	1	US-07-911-531-18	Sequence 18, Appl
C 114	13.8	76.7	3541	4	US-09-963-137-160	Sequence 160, App	C 187	13.4	74.4	2487	3	US-08-481-435-1	Sequence 1, Appl
C 115	13.8	76.7	3541	4	US-09-023-655-1205	Sequence 1205, App	C 188	13.4	74.4	2487	4	US-09-543-681A-3399	Sequence 3999, Ap
C 116	13.8	76.7	3794	4	US-09-710-279-3705	Sequence 3705, Ap	C 189	13.4	74.4	2710	4	US-08-956-171E-421	Sequence 421, App
C 117	13.8	76.7	3969	1	US-09-248-796A-5440	Sequence 5440, Ap	C 190	13.4	74.4	2710	4	US-08-956-171E-421	Sequence 421, App
C 118	13.8	76.7	4234	1	US-08-445-038B-1	Sequence 1, Appl	C 191	13.4	74.4	3177	4	US-09-710-279-9335	Sequence 3359, Ap
C 119	13.8	76.7	4234	1	US-08-445-038B-1	Sequence 1, Appl	C 192	13.4	74.4	3177	4	US-09-710-279-9335	Sequence 3359, Ap
C 120	13.8	76.7	4234	1	US-08-805-445-1	Sequence 1, Appl	C 193	13.4	74.4	3828	4	US-09-710-279-93921	Sequence 3921, Ap
C 121	13.8	76.7	4234	2	US-08-064-067D-1	Sequence 1, Appl	C 194	13.4	74.4	4302	4	US-08-956-171E-276	Sequence 276, Appl
C 122	13.8	76.7	4234	2	US-09-066-208-1	Sequence 1, Appl	C 195	13.4	74.4	4846	4	US-08-956-171E-276	Sequence 276, Appl
C 123	13.8	76.7	4748	3	US-09-331-581-1	Sequence 1, Appl	C 196	13.4	74.4	4846	4	US-08-956-171E-276	Sequence 276, Appl
C 124	13.8	76.7	6256	2	US-08-475-991A-1	Sequence 1, Appl	C 197	13.4	74.4	5537	1	US-08-135-511-32	Sequence 32, Appl
C 125	13.8	76.7	6256	2	US-08-567-375-1	Sequence 1, Appl	C 198	13.4	74.4	5537	1	US-08-483-852-9	Sequence 9, Appl
C 126	13.8	76.7	6256	2	US-08-587-680A-1	Sequence 1, Appl	C 199	13.4	74.4	5537	1	US-08-361-458-4	Sequence 4, Appl
C 127	13.8	76.7	7286	3	US-09-331-581-3	Sequence 3, Appl	C 200	13.4	74.4	5537	1		

C 247	13.2	73.3	620	4	US-09-270-767-25488	Sequence 28488, A	320	13.2	73.3	2521	1	US-08-332-638-59	Sequence 59, Appl
C 248	13.2	73.3	624	1	US-09-540-236-145	Sequence 145, App	321	13.2	73.3	2525	1	US-08-294-312B-1	Sequence 1, Appl1
C 249	13.2	73.3	662	1	US-07-807-043B-10	Sequence 10, Appl	322	13.2	73.3	2525	4	US-08-468-024B-1	Sequence 1, Appl1
C 250	13.2	73.3	662	1	US-08-299-849B-10	Sequence 10, Appl	323	13.2	73.3	2525	4	US-08-187-757D-1	Sequence 1, Appl1
C 251	13.2	73.3	662	2	US-08-142-168A-10	Sequence 10, Appl	324	13.2	73.3	2525	4	US-08-465-679-1	Sequence 1, Appl1
C 252	13.2	73.3	662	3	US-08-967-727-10	Sequence 10, Appl	325	13.2	73.3	2525	4	US-08-210-143C-1	Sequence 1, Appl1
C 253	13.2	73.3	662	3	US-08-037-230D-10	Sequence 10, Appl	326	13.2	73.3	2580	2	US-08-887-798-1	Sequence 1, Appl1
C 254	13.2	73.3	662	4	US-09-583-850-10	Sequence 10, Appl	327	13.2	73.3	2908	4	US-09-930-181-1	Sequence 1, Appl1
C 255	13.2	73.3	662	4	US-09-579-197-10	Sequence 10, Appl	328	13.2	73.3	2943	1	US-08-042-747A-7	Sequence 7, Appl1
C 256	13.2	73.3	662	4	US-09-404-026-10	Sequence 10, Appl	329	13.2	73.3	2980	4	US-10-003-690-1	Sequence 1, Appl1
C 257	13.2	73.3	662	4	US-09-312-464-10	Sequence 10, Appl	330	13.2	73.3	2982	4	US-09-710-279-385	Sequence 385, App
C 258	13.2	73.3	662	4	US-09-583-110-960	Sequence 960, App	331	13.2	73.3	3045	4	US-09-633-238-10	Sequence 10, Appl
C 259	13.2	73.3	684	4	US-09-114-000C-2631	Sequence 632, App	332	13.2	73.3	3120	4	US-09-432-468A-14	Sequence 4, Appl1
C 260	13.2	73.3	705	3	US-08-998-416-632	Sequence 2631, App	333	13.2	73.3	3177	1	US-08-042-747A-4	Sequence 4, Appl1
C 261	13.2	73.3	815	4	US-08-482-709A-5	Sequence 5, Appl1	334	13.2	73.3	3277	4	US-09-293-549-5	Sequence 3, Appl1
C 262	13.2	73.3	815	4	US-09-248-796A-2554	Sequence 2554, App	335	13.2	73.3	3364	4	US-09-930-181-3	Sequence 3, Appl1
C 263	13.2	73.3	898	4	US-09-270-767-13110	Sequence 13110, A	336	13.2	73.3	3472	6	5244792-1	Parent No. 5244792
C 264	13.2	73.3	909	1	US-08-701-952A-3	Sequence 3, Appl1	337	13.2	73.3	3494	4	US-09-809-665A-84	Sequence 84, Appl
C 265	13.2	73.3	909	1	US-08-938-471-3	Sequence 3, Appl1	338	13.2	73.3	3821	4	US-08-956-171E-304	Sequence 304, App
C 266	13.2	73.3	957	4	US-09-543-681A-2212	Sequence 2212, App	339	13.2	73.3	3821	4	US-08-781-986A-304	Sequence 880, App
C 267	13.2	73.3	983	4	US-09-489-039A-33737	Sequence 3337, App	340	13.2	73.3	4017	4	US-10-101-464A-880	Sequence 830, App
C 268	13.2	73.3	1000	4	US-09-270-767-73268	Sequence 7326, App	341	13.2	73.3	4056	4	US-10-101-464A-833	Sequence 3879, App
C 269	13.2	73.3	1000	4	US-09-270-767-72608	Sequence 22608, A	342	13.2	73.3	4158	4	US-09-710-279-3879	Sequence 12126, A
C 270	13.2	73.3	1017	3	US-09-330-611-5	Sequence 5, Appl1	343	13.2	73.3	4484	4	US-09-252-991A-12126	Sequence 12392, A
C 271	13.2	73.3	1050	4	US-09-252-991A-12982	Sequence 12982, A	344	13.2	73.3	4884	4	US-09-252-991A-12126	Sequence 6, Appl1
C 272	13.2	73.3	1092	4	US-09-252-991A-13142	Sequence 13142, A	345	13.2	73.3	5339	3	US-09-073-354-6	Sequence 6, Appl1
C 273	13.2	73.3	1146	4	US-09-710-279-3055	Sequence 3055, App	346	13.2	73.3	5339	3	US-08-656-005A-6	Sequence 6, Appl1
C 274	13.2	73.3	1182	4	US-09-252-991A-12239	Sequence 12239, App	347	13.2	73.3	5339	3	US-09-073-285-6	Sequence 6, Appl1
C 275	13.2	73.3	1197	4	US-08-489-039A-6585	Sequence 6585, App	348	13.2	73.3	5339	3	US-09-363-095-6	Sequence 6, Appl1
C 276	13.2	73.3	1344	4	US-09-252-991A-6450	Sequence 6450, App	349	13.2	73.3	5342	3	US-09-418-027-6	Sequence 1, Appl1
C 277	13.2	73.3	1366	4	US-09-270-767-12957	Sequence 12957, A	350	13.2	73.3	5342	3	US-08-902-632-1	Sequence 5, Appl1
C 278	13.2	73.3	1380	1	US-07-936-163-1	Sequence 1, Appl1	351	13.2	73.3	5342	3	US-09-073-354-5	Sequence 5, Appl1
C 279	13.2	73.3	1440	4	US-09-489-039A-6807	Sequence 6807, App	352	13.2	73.3	5342	3	US-08-656-005A-5	Sequence 5, Appl1
C 280	13.2	73.3	1506	4	US-09-252-991A-6098	Sequence 6098, App	353	13.2	73.3	5342	3	US-09-073-285-6	Sequence 6, Appl1
C 281	13.2	73.3	1560	3	US-09-105-039A-1	Sequence 1, Appl1	354	13.2	73.3	5342	3	US-09-363-095-5	Sequence 6, Appl1
C 282	13.2	73.3	1583	4	US-09-620-312D-974	Sequence 974, App	355	13.2	73.3	5342	3	US-09-418-027-5	Sequence 5, Appl1
C 283	13.2	73.3	1586	3	US-09-352-772-3	Sequence 3, Appl1	356	13.2	73.3	5883	1	US-08-404-445-1	Sequence 1, Appl1
C 284	13.2	73.3	1589	3	US-09-270-767-13252	Sequence 13252, A	357	13.2	73.3	7060	4	US-09-221-011B-479	Sequence 479, App
C 285	13.2	73.3	1593	4	US-09-134-000C-1665	Sequence 1665, App	358	13.2	73.3	7762	4	US-09-220-123-24	Sequence 24, Appl
C 286	13.2	73.3	1599	4	US-09-252-991A-11471	Sequence 11471, A	359	13.2	73.3	7964	4	US-09-521-511C-12	Sequence 12, Appl
C 287	13.2	73.3	1605	3	US-09-105-039A-3	Sequence 3, Appl1	360	13.2	73.3	13636	3	US-08-961-527-171	Sequence 171, App
C 288	13.2	73.3	1607	4	US-09-081-087-9	Sequence 9, Appl1	361	13.2	73.3	13710	4	US-08-976-259-70	Sequence 70, Appl
C 289	13.2	73.3	1674	4	US-09-543-681A-2434	Sequence 2434, App	362	13.2	73.3	17710	3	US-09-956-004-70	Sequence 18, Appl
C 290	13.2	73.3	1740	4	US-09-252-991A-12801	Sequence 12801, A	363	13.2	73.3	33000	3	US-09-215-694-18	Sequence 10, Appl
C 291	13.2	73.3	1752	3	US-09-188-930-225	Sequence 225, App	364	13.2	73.3	33000	3	US-09-567-666-10	Sequence 31, Appl
C 292	13.2	73.3	1752	3	US-09-312-283C-225	Sequence 225, App	365	13.2	73.3	34945	4	US-09-596-002-31	Sequence 32, Appl
C 293	13.2	73.3	1770	4	US-09-252-991A-12285	Sequence 12285, A	366	13.2	73.3	34945	4	US-09-497-855A-32	Sequence 11, Appl
C 294	13.2	73.3	1770	4	US-09-252-991A-12418	Sequence 12418, A	367	13.2	73.3	36792	4	US-10-172-911-11	Sequence 20, Appl
C 295	13.2	73.3	1797	4	US-09-774-528-221	Sequence 221, App	368	13.2	73.3	37000	4	US-08-724-394A-20	Sequence 21, Appl
C 296	13.2	73.3	1830	4	US-09-557-921-1	Sequence 1, Appl1	369	13.2	73.3	246240	2	US-08-724-394A-21	Sequence 22, Appl
C 297	13.2	73.3	1830	4	US-09-630-983A-2	Sequence 2, Appl1	370	13.2	73.3	246240	2	US-08-916-421B-1	Sequence 1, Appl1
C 298	13.2	73.3	2000	4	US-09-516-052-23	Sequence 23, Appl1	371	13.2	73.3	246240	2	US-09-692-570-1	Sequence 1, Appl1
C 299	13.2	73.3	2015	4	US-09-774-528-398	Sequence 398, App	372	13.2	73.3	1664976	4	US-09-134-001C-2386	Sequence 2386, App
C 300	13.2	73.3	2028	4	US-09-270-767-12056	Sequence 12056, App	373	13.2	73.3	204	3	US-09-513-294A-1202	Sequence 55, Appl1
C 301	13.2	73.3	2077	1	US-08-217-337-7	Sequence 7, Appl1	374	13.2	73.3	284	4	US-08-821-994-55	Sequence 24, Appl1
C 302	13.2	73.3	2112	4	US-10-116-326-5	Sequence 5, Appl1	375	13.2	73.3	322	3	US-09-556-877-24	Sequence 24, Appl1
C 303	13.2	73.3	2114	4	US-09-248-796A-216	Sequence 216, App	376	13.2	73.3	363	4	US-09-556-877-24	Sequence 24, Appl1
C 304	13.2	73.3	2249	4	US-09-270-767-12680	Sequence 12680, A	377	13.2	73.3	363	4	US-09-288-594A-24	Sequence 24, Appl1
C 305	13.2	73.3	2289	4	US-10-116-326-3	Sequence 3, Appl1	378	13.2	73.3	363	4	US-09-620-412C-24	Sequence 24, Appl1
C 306	13.2	73.3	2334	4	US-09-603-787A-181	Sequence 181, App	379	13.2	73.3	363	4	US-09-620-412C-24	Sequence 24, Appl1
C 307	13.2	73.3	2334	4	US-10-003-690-3	Sequence 3, Appl1	380	13.2	73.3	363	4	US-09-620-412C-24	Sequence 24, Appl1
C 308	13.2	73.3	2337	4	US-10-116-326-1	Sequence 1, Appl1	381	13.2	73.3	363	4	US-09-620-412C-24	Sequence 24, Appl1
C 309	13.2	73.3	2364	4	US-09-252-991A-12368	Sequence 12368, A	382	13.2	73.3	400	4	US-08-956-171E-1976	Sequence 1976, App
C 310	13.2	73.3	2385	4	US-08-352-902D-145	Sequence 145, App	383	13.2	73.3	400	4	US-08-781-986A-1976	Sequence 1976, App
C 311	13.2	73.3	2484	2	US-08-203-521-8	Sequence 8, Appl1	384	13.2	73.3	432	4	US-09-107-532A-2097	Sequence 2097, App
C 312	13.2	73.3	2484	3	US-08-961-810-4	Sequence 4, Appl1	385	13.2	73.3	432	4	US-09-248-796A-2599	Sequence 2599, App
C 313	13.2	73.3	2484	3	US-08-352-902D-4	Sequence 4, Appl1	386	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 314	13.2	73.3	2484	3	US-08-265-503B-4	Sequence 4, Appl1	387	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 315	13.2	73.3	2484	4	US-09-708-200-16	Sequence 16, Appl1	388	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 316	13.2	73.3	2484	4	US-09-788-657-10	Sequence 10, Appl	389	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 317	13.2	73.3	2484	4	US-09-712-691-14	Sequence 14, Appl	390	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 318	13.2	73.3	2521	1	US-08-188-228-59	Sequence 59, Appl	391	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App
C 319	13.2	73.3	2521	1	US-08-332-644-53	Sequence 53, Appl	392	13.2	73.3	432	4	US-09-513-999E-8187	Sequence 8187, App

C 333	13	72.2	550	3	US-09-127-480-110	Sequence 110, App	466	13	72.2	2509	3	US-08-229-050-14	Sequence 14, Appl
C 334	13	72.2	550	3	US-08-496-841C-110	Sequence 110, App	467	13	72.2	2509	3	US-08-801-563-14	Sequence 14, Appl
C 335	13	72.2	550	3	US-09-124-523-110	Sequence 110, App	468	13	72.2	2509	3	US-09-710-279-4170	Sequence 4170, Ap
C 336	13	72.2	550	4	US-09-636-796A-110	Sequence 110, App	469	13	72.2	2509	4	US-09-710-279-4151	Sequence 4151, Ap
C 337	13	72.2	550	4	US-08-431-048F-110	Sequence 110, App	470	13	72.2	3073	4	US-09-710-279-3606	Sequence 3606, Ap
C 338	13	72.2	611	4	US-09-376-113-4	Sequence 4, Appl	471	13	72.2	3132	4	US-09-710-279-3455	Sequence 3455, Ap
C 339	13	72.2	631	3	US-09-449-218D-9	Sequence 6, Appl	472	13	72.2	3371	4	US-09-710-279-3490	Sequence 3490, Ap
C 400	13	72.2	642	3	US-09-449-218D-45	Sequence 9, Appl	473	13	72.2	3465	4	US-09-710-279-1779	Sequence 1779, Ap
C 401	13	72.2	642	4	US-09-668-529A-9	Sequence 45, Appl	474	13	72.2	3543	4	US-09-252-991A-15893	Sequence 1593, A
C 402	13	72.2	642	4	US-09-668-529A-45	Sequence 45, Appl	475	13	72.2	3722	4	US-09-910-039-98	Sequence 98, Appl
C 403	13	72.2	642	4	US-09-668-529A-9	Sequence 45, Appl	476	13	72.2	3744	2	US-08-348-353-16	Sequence 16, Appl
C 404	13	72.2	642	4	US-09-668-529A-45	Sequence 45, Appl	477	13	72.2	3744	2	US-08-465-965-16	Sequence 16, Appl
C 405	13	72.2	642	4	US-09-668-529A-45	Sequence 45, Appl	478	13	72.2	3744	2	US-08-465-965-16	Sequence 16, Appl
C 406	13	72.2	679	3	US-08-821-994-54	Sequence 54, Appl	479	13	72.2	4103	4	US-09-710-279-3532	Sequence 3532, Ap
C 407	13	72.2	690	3	US-08-821-994-53	Sequence 53, Appl	480	13	72.2	4257	4	US-09-252-991A-15796	Sequence 15796, A
C 408	13	72.2	729	3	US-08-858-207A-198	Sequence 198, App	481	13	72.2	4483	1	US-08-181-211A-5	Sequence 5, Appl
C 409	13	72.2	855	3	US-09-248-796A-6820	Sequence 6820, App	482	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 410	13	72.2	879	4	US-09-583-110-132	Sequence 132, App	483	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 411	13	72.2	888	4	US-09-710-279-2945	Sequence 2945, App	484	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 412	13	72.2	1102	3	US-08-821-994-86	Sequence 86, Appl	485	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 413	13	72.2	1125	4	US-09-328-352-2336	Sequence 2336, App	486	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 414	13	72.2	1197	4	US-09-614-221A-549	Sequence 549, App	487	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 415	13	72.2	1256	4	US-09-556-877-21	Sequence 21, Appl	488	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 416	13	72.2	1256	4	US-09-288-594A-21	Sequence 21, Appl	489	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 417	13	72.2	1256	4	US-09-620-412C-21	Sequence 21, Appl	490	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 418	13	72.2	1256	4	US-09-410-568-21	Sequence 21, Appl	491	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 419	13	72.2	1256	4	US-09-598-419-21	Sequence 21, Appl	492	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 420	13	72.2	1256	4	US-08-181-211A-20	Sequence 20, Appl	493	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 421	13	72.2	1358	1	US-08-449-315-20	Sequence 20, Appl	494	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 422	13	72.2	1358	1	US-08-444-803-20	Sequence 20, Appl	495	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 423	13	72.2	1358	1	US-08-449-043-20	Sequence 20, Appl	496	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 424	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	497	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 425	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	498	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 426	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	499	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 427	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	500	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 428	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	501	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 429	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	502	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 430	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	503	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 431	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	504	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 432	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	505	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 433	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	506	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 434	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	507	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 435	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	508	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 436	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	509	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 437	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	510	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 438	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	511	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 439	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	512	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 440	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	513	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 441	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	514	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 442	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	515	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 443	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	516	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 444	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	517	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 445	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	518	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 446	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	519	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 447	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	520	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 448	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	521	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 449	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	522	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 450	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	523	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 451	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	524	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 452	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	525	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 453	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	526	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 454	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	527	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 455	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	528	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 456	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	529	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 457	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	530	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 458	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	531	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 459	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	532	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 460	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	533	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 461	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	534	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 462	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	535	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 463	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	536	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 464	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	537	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl
C 465	13	72.2	1358	2	US-08-456-265A-20	Sequence 20, Appl	538	13	72.2	4483	2	US-08-444-803-5	Sequence 5, Appl

539	12.8	71.1	144	4	US-09-313-294A-2557	Sequence 2557, Ap	c 612	12.8	71.1	631	4	US-09-270-767-10760	Sequence 10760, A
540	12.8	71.1	188	4	US-09-248-796A-10752	Sequence 10752, A	c 613	12.8	71.1	631	4	US-09-270-767-15278	Sequence 15278, A
541	12.8	71.1	200	4	US-09-313-294A-3705	Sequence 3705, Ap	c 614	12.8	71.1	654	4	US-09-583-110-150	Sequence 150, App
542	12.8	71.1	240	4	US-09-513-999C-33512	Sequence 33512, A	c 615	12.8	71.1	654	4	US-09-583-110-751	Sequence 751, App
543	12.8	71.1	244	5	US-08-238-821B-47	Sequence 47, Appl	c 616	12.8	71.1	654	4	US-09-583-110-1831	Sequence 1831, Ap
544	12.8	71.1	244	5	PCT-US95-05744-47	Sequence 47, Appl	c 617	12.8	71.1	670	4	US-09-270-767-10084	Sequence 10084, A
545	12.8	71.1	248	4	US-09-016-434-524	Sequence 524, App	c 618	12.8	71.1	690	2	US-08-612-840A-7	Sequence 7, Appl1
546	12.8	71.1	250	4	US-09-270-767-8208	Sequence 8208, Ap	c 619	12.8	71.1	690	2	US-09-969-763-2	Sequence 2, Appl1
547	12.8	71.1	250	4	US-09-270-767-23490	Sequence 23490, A	c 620	12.8	71.1	707	1	US-08-351-176-1	Sequence 1, Appl1
548	12.8	71.1	258	4	US-09-248-796A-11874	Sequence 11874, A	c 621	12.8	71.1	707	1	US-09-455-976-1	Sequence 1, Appl1
549	12.8	71.1	274	4	US-09-270-767-1077	Sequence 1077, Ap	c 622	12.8	71.1	707	3	US-09-222-227-1	Sequence 1, Appl1
550	12.8	71.1	274	4	US-09-270-767-16359	Sequence 16359, A	c 623	12.8	71.1	712	3	US-08-998-416-867	Sequence 867, App
551	12.8	71.1	274	4	US-09-513-999C-23822	Sequence 23822, A	c 624	12.8	71.1	723	3	US-08-998-416-635	Sequence 635, App
552	12.8	71.1	275	4	US-09-313-294A-3687	Sequence 3687, Ap	c 625	12.8	71.1	724	4	US-09-533-008-40	Sequence 40, Appl
553	12.8	71.1	278	4	US-09-313-294A-5760	Sequence 5760, Ap	c 626	12.8	71.1	729	4	US-09-634-238-133	Sequence 133, App
554	12.8	71.1	282	4	US-09-248-796A-1188	Sequence 1188, Ap	c 627	12.8	71.1	729	4	US-09-270-767-1796	Sequence 1796, Ap
555	12.8	71.1	284	5	US-08-238-821B-45	Sequence 45, Appl	c 628	12.8	71.1	729	4	US-09-270-767-17078	Sequence 17078, A
556	12.8	71.1	284	5	PCT-US95-05744-45	Sequence 45, Appl	c 629	12.8	71.1	759	4	US-09-523-029-85	Sequence 95, Appl
557	12.8	71.1	286	4	US-09-313-294A-6838	Sequence 6838, Ap	c 630	12.8	71.1	776	4	US-09-270-767-3669	Sequence 3669, Ap
558	12.8	71.1	288	4	US-09-313-294A-6654	Sequence 6654, Ap	c 631	12.8	71.1	776	4	US-09-270-767-18951	Sequence 18951, A
559	12.8	71.1	309	4	US-08-916-442-19	Sequence 19, Appl	c 632	12.8	71.1	784	3	US-09-004-838-64	Sequence 64, Appl
560	12.8	71.1	309	4	US-09-731-924A-19	Sequence 19, Appl	c 633	12.8	71.1	789	4	US-09-248-796A-2549	Sequence 2549, Ap
561	12.8	71.1	309	4	US-09-317-641-19	Sequence 19, Appl	c 634	12.8	71.1	792	4	US-09-252-991A-7444	Sequence 7444, A
562	12.8	71.1	373	4	US-09-404-879A-190	Sequence 190, App	c 635	12.8	71.1	825	4	US-09-248-796A-10866	Sequence 10866, A
563	12.8	71.1	373	4	US-09-338-933-190	Sequence 190, App	c 636	12.8	71.1	826	2	US-08-238-821B-51	Sequence 51, Appl
564	12.8	71.1	373	4	US-09-215-681-190	Sequence 190, App	c 637	12.8	71.1	826	2	US-08-238-821B-61	Sequence 61, Appl
565	12.8	71.1	373	4	US-09-216-003A-190	Sequence 190, App	c 638	12.8	71.1	826	5	PCT-US95-05744-51	Sequence 51, Appl
566	12.8	71.1	373	4	US-09-667-857-190	Sequence 190, App	c 639	12.8	71.1	826	5	PCT-US95-05744-51	Sequence 61, Appl
567	12.8	71.1	373	4	US-09-404-879A-184	Sequence 184, App	c 640	12.8	71.1	864	4	US-09-248-796A-567	Sequence 567, App
568	12.8	71.1	375	4	US-09-338-933-184	Sequence 184, App	c 641	12.8	71.1	872	4	US-09-110-532A-87	Sequence 87, Appl
569	12.8	71.1	375	4	US-09-215-681-184	Sequence 184, App	c 642	12.8	71.1	882	4	US-09-710-279-1633	Sequence 1633, Ap
570	12.8	71.1	375	4	US-09-216-003A-184	Sequence 184, App	c 643	12.8	71.1	887	4	US-09-221-017B-701	Sequence 701, App
571	12.8	71.1	375	4	US-09-667-857-184	Sequence 184, App	c 644	12.8	71.1	888	3	US-08-801-740-4	Sequence 4, Appl1
572	12.8	71.1	393	4	US-09-328-352-1392	Sequence 1392, App	c 645	12.8	71.1	892	3	US-08-801-740-4	Sequence 4, Appl1
573	12.8	71.1	405	4	US-09-248-796A-12580	Sequence 12580, A	c 646	12.8	71.1	892	3	US-09-171-209-5	Sequence 5, Appl1
574	12.8	71.1	414	4	US-09-710-279-529	Sequence 529, App	c 647	12.8	71.1	896	4	US-09-252-991A-1504	Sequence 1504, Ap
575	12.8	71.1	414	4	US-09-270-767-11910	Sequence 11910, A	c 648	12.8	71.1	936	4	US-09-107-532A-1365	Sequence 1365, Ap
576	12.8	71.1	462	4	US-08-916-442-20	Sequence 20, Appl	c 649	12.8	71.1	951	4	US-09-252-991A-1256	Sequence 1256, Ap
577	12.8	71.1	462	4	US-09-731-924A-20	Sequence 20, Appl	c 650	12.8	71.1	966	4	US-09-252-991A-7256	Sequence 7256, Ap
578	12.8	71.1	462	4	US-09-317-641-20	Sequence 20, Appl	c 651	12.8	71.1	1007	4	US-09-270-767-23881	Sequence 23881, A
579	12.8	71.1	462	4	US-09-270-767-11694	Sequence 11694, A	c 652	12.8	71.1	1017	4	US-09-252-991A-10799	Sequence 10799, A
580	12.8	71.1	463	4	US-09-270-767-14424	Sequence 14424, A	c 653	12.8	71.1	1017	4	US-09-690-454-35	Sequence 35, Appl
581	12.8	71.1	468	4	US-09-621-976-17544	Sequence 11, Appl	c 654	12.8	71.1	1032	4	US-09-328-352-1642	Sequence 1642, Ap
582	12.8	71.1	472	4	US-09-702-705-11	Sequence 11, Appl	c 655	12.8	71.1	1032	4	US-09-134-000C-2551	Sequence 349, App
583	12.8	71.1	472	4	US-09-736-457-11	Sequence 11, Appl	c 656	12.8	71.1	1038	4	US-08-599-171A-23	Sequence 23, Appl
584	12.8	71.1	472	4	US-09-614-124B-11	Sequence 11, Appl	c 657	12.8	71.1	1050	1	US-08-646-590B-23	Sequence 23, Appl
585	12.8	71.1	472	4	US-09-671-325-11	Sequence 11, Appl	c 658	12.8	71.1	1065	2	US-09-069-226-23	Sequence 23, Appl
586	12.8	71.1	472	4	US-09-589-184-11	Sequence 11, Appl	c 659	12.8	71.1	1065	3	US-09-412-184-23	Sequence 23, Appl
587	12.8	71.1	472	4	US-09-658-824-11	Sequence 11, Appl	c 660	12.8	71.1	1065	3	US-09-412-184-23	Sequence 33, App
588	12.8	71.1	472	4	US-09-916-442-15	Sequence 15, Appl	c 661	12.8	71.1	1065	3	US-09-248-796A-10719	Sequence 10719, A
589	12.8	71.1	520	4	US-09-317-641-15	Sequence 15, Appl	c 662	12.8	71.1	1065	4	US-09-252-991A-10448	Sequence 10448, A
590	12.8	71.1	520	4	US-08-916-442-12	Sequence 12, Appl	c 663	12.8	71.1	1065	4	US-09-328-352-374	Sequence 374, App
591	12.8	71.1	530	4	US-09-916-442-12	Sequence 12, Appl	c 664	12.8	71.1	1065	4	US-09-328-352-374	Sequence 374, App
592	12.8	71.1	530	4	US-09-731-924A-12	Sequence 12, Appl	c 665	12.8	71.1	1101	4	US-09-134-000C-3188	Sequence 3188, Ap
593	12.8	71.1	530	4	US-09-317-641-12	Sequence 12, Appl	c 666	12.8	71.1	1107	4	US-09-252-991A-9656	Sequence 9656, Ap
594	12.8	71.1	538	4	US-09-270-767-705	Sequence 705, App	c 667	12.8	71.1	1116	4	US-09-328-352-374	Sequence 374, App
595	12.8	71.1	538	4	US-09-270-767-15987	Sequence 15987, A	c 668	12.8	71.1	1116	4	US-09-328-352-374	Sequence 374, App
596	12.8	71.1	544	4	US-10-101-464A-189	Sequence 189, App	c 669	12.8	71.1	1173	4	US-09-328-352-374	Sequence 374, App
597	12.8	71.1	567	4	US-09-621-976-16902	Sequence 16902, A	c 670	12.8	71.1	1173	4	US-09-448-796A-1740	Sequence 1740, Ap
598	12.8	71.1	576	4	US-09-270-767-14218	Sequence 14218, A	c 671	12.8	71.1	1182	4	US-09-448-796A-1740	Sequence 1740, Ap
599	12.8	71.1	586	4	US-08-779-764A-2	Sequence 2, Appl1	c 672	12.8	71.1	1200	4	US-09-602-787A-465	Sequence 465, App
600	12.8	71.1	603	3	US-08-779-764A-3	Sequence 3, Appl1	c 673	12.8	71.1	1200	4	US-09-602-787A-465	Sequence 9666, Ap
601	12.8	71.1	603	3	US-08-779-764A-4	Sequence 4, Appl1	c 674	12.8	71.1	1239	4	US-09-731-924A-1	Sequence 1, Appl1
602	12.8	71.1	603	3	US-08-779-764A-5	Sequence 5, Appl1	c 675	12.8	71.1	1244	4	US-08-916-442-1	Sequence 1, Appl1
603	12.8	71.1	603	3	US-09-563-456-2	Sequence 2, Appl1	c 676	12.8	71.1	1244	4	US-08-916-442-1	Sequence 1, Appl1
604	12.8	71.1	603	3	US-09-563-456-3	Sequence 3, Appl1	c 677	12.8	71.1	1250	4	US-09-317-641-1	Sequence 37, Appl
605	12.8	71.1	603	3	US-09-563-456-4	Sequence 4, Appl1	c 678	12.8	71.1	1250	4	US-09-317-641-1	Sequence 37, Appl
606	12.8	71.1	603	3	US-09-563-456-5	Sequence 5, Appl1	c 679	12.8	71.1	1260	4	US-09-448-796A-13993	Sequence 3293, Ap
607	12.8	71.1	603	3	US-08-369-822C-3	Sequence 3, Appl1	c 680	12.8	71.1	1271	4	US-09-270-767-11058	Sequence 11058, A
608	12.8	71.1	609	3	US-08-582-776C-3	Sequence 3, Appl1	c 681	12.8	71.1	1272	4	US-09-248-796A-1740	Sequence 1740, Ap
609	12.8	71.1	609	3	US-08-434-831B-3	Sequence 3, Appl1	c 682	12.8	71.1	1272	4	US-09-248-796A-1740	Sequence 1740, Ap
610	12.8	71.1	612	4	US-09-252-991A-1623	Sequence 1623, Ap	c 683	12.8	71.1	1272	4	US-09-187-999-30	Sequence 30, Appl
611	12.8	71.1	612	4	US-09-252-991A-1623	Sequence 1623, Ap	c 684	12.8	71.1	1284	4	US-09-540-236-26	Sequence 26, Appl

C 685	12.8	71.1	1296	4	US-09-252-991A-1670	Sequence 1670, Ap	C 758	12.8	71.1	2694	1	US-08-465-995A-3	Sequence 3, Appl1
C 686	12.8	71.1	1323	4	US-09-489-039A-5432	Sequence 5432, Ap	C 759	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 687	12.8	71.1	1371	4	US-09-350-756-3	Sequence 3, Appl1	C 760	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 688	12.8	71.1	1374	4	US-09-123-764A-3	Sequence 2, Appl1	C 761	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 689	12.8	71.1	1405	4	US-09-270-767-10838	Sequence 10838, A	C 762	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 690	12.8	71.1	1449	4	US-09-252-991A-6087	Sequence 6087, Ap	C 763	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 691	12.8	71.1	1473	4	US-08-277-031B-15	Sequence 15, Appl1	C 764	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 692	12.8	71.1	1497	4	US-09-248-796A-11414	Sequence 11414, A	C 765	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 693	12.8	71.1	1500	4	US-09-023-778-1	Sequence 1, Appl1	C 766	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 694	12.8	71.1	1509	4	US-09-248-796A-2721	Sequence 2721, Ap	C 767	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 695	12.8	71.1	1530	4	US-09-328-352-739	Sequence 739, Ap	C 768	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 696	12.8	71.1	1537	4	US-09-270-767-13829	Sequence 13829, A	C 769	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 697	12.8	71.1	1594	4	US-09-720-767-14907	Sequence 14907, A	C 770	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 698	12.8	71.1	1604	4	US-07-621-670-2	Sequence 2, Appl1	C 771	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 699	12.8	71.1	1607	4	US-08-921-655-178	Sequence 178, Ap	C 772	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 700	12.8	71.1	1659	4	US-08-941-532-5	Sequence 5, Appl1	C 773	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 701	12.8	71.1	1659	4	US-09-270-767-12870	Sequence 12870, A	C 774	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 702	12.8	71.1	1698	4	US-09-248-796A-5105	Sequence 5105, Ap	C 775	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 703	12.8	71.1	1716	3	US-09-049-681A-3	Sequence 3, Appl1	C 776	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 704	12.8	71.1	1716	3	US-09-497-897-3	Sequence 3, Appl1	C 777	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 705	12.8	71.1	1746	2	US-08-201-118-2	Sequence 2, Appl1	C 778	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 706	12.8	71.1	1746	2	US-08-238-821B-2	Sequence 2, Appl1	C 779	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 707	12.8	71.1	1746	2	US-09-107-532A-3181	Sequence 3181, Ap	C 780	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 708	12.8	71.1	1746	2	US-09-023-655-1059	Sequence 1059, Ap	C 781	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 709	12.8	71.1	1746	5	PCT-US95-05744-2	Sequence 2, Appl1	C 782	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 710	12.8	71.1	1765	4	US-09-559-035-11	Sequence 11, Appl1	C 783	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 711	12.8	71.1	1766	4	US-09-270-767-10946	Sequence 10946, A	C 784	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 712	12.8	71.1	1816	4	US-09-976-594-838	Sequence 838, Ap	C 785	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 713	12.8	71.1	1827	4	US-09-270-767-14852	Sequence 14852, A	C 786	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 714	12.8	71.1	1834	4	US-10-101-464A-262	Sequence 262, Ap	C 787	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 715	12.8	71.1	1870	2	US-08-359-705B-3	Sequence 3, Appl1	C 788	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 716	12.8	71.1	1870	2	US-08-286-846A-3	Sequence 3, Appl1	C 789	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 717	12.8	71.1	1870	2	US-08-457-860A-3	Sequence 3, Appl1	C 790	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 718	12.8	71.1	1870	2	US-08-444-622A-3	Sequence 3, Appl1	C 791	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 719	12.8	71.1	1870	3	US-08-942-562-3	Sequence 3, Appl1	C 792	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 720	12.8	71.1	1870	3	US-09-156-923-3	Sequence 3, Appl1	C 793	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 721	12.8	71.1	1920	3	US-09-509-814A-3	Sequence 3, Appl1	C 794	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 722	12.8	71.1	1920	3	US-09-134-001C-2824	Sequence 2824, Ap	C 795	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 723	12.8	71.1	1920	4	US-09-920-954-3	Sequence 3, Appl1	C 796	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 724	12.8	71.1	1935	4	US-09-270-767-11338	Sequence 11338, A	C 797	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 725	12.8	71.1	1935	4	US-09-248-796A-6538	Sequence 6538, Ap	C 798	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 726	12.8	71.1	1950	4	US-09-727-628-1	Sequence 1, Appl1	C 799	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 727	12.8	71.1	1973	4	US-09-270-767-5756	Sequence 5756, Ap	C 800	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 728	12.8	71.1	1973	4	US-09-620-312D-715	Sequence 21038, A	C 801	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 729	12.8	71.1	1981	4	US-09-270-767-21038	Sequence 715, Ap	C 802	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 730	12.8	71.1	2019	4	US-09-270-767-14598	Sequence 14598, A	C 803	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 731	12.8	71.1	2061	4	US-08-252-991A-1544	Sequence 1544, A	C 804	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 732	12.8	71.1	2092	4	US-09-458-253B-13	Sequence 13, Appl1	C 805	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 733	12.8	71.1	2100	4	US-09-270-767-11711	Sequence 11711, A	C 806	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 734	12.8	71.1	2106	1	US-07-718-535-4	Sequence 4, Appl1	C 807	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 735	12.8	71.1	2106	1	US-08-161-999-4	Sequence 38, Appl1	C 808	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 736	12.8	71.1	2122	4	US-09-367-895-38	Sequence 117, Ap	C 809	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 737	12.8	71.1	2190	2	US-09-023-655-117	Sequence 1, Appl1	C 810	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 738	12.8	71.1	2268	2	US-08-880-853-1	Sequence 1, Appl1	C 811	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 739	12.8	71.1	2268	2	US-09-099-125A-1	Sequence 1, Appl1	C 812	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 740	12.8	71.1	2268	2	US-09-099-125A-1	Sequence 1, Appl1	C 813	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 741	12.8	71.1	2268	2	US-09-197-008-1	Sequence 1, Appl1	C 814	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 742	12.8	71.1	2268	2	US-09-032-476-1	Sequence 1, Appl1	C 815	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 743	12.8	71.1	2268	3	US-08-890-854-1	Sequence 1, Appl1	C 816	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 744	12.8	71.1	2268	3	US-09-023-324-1	Sequence 1, Appl1	C 817	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 745	12.8	71.1	2268	4	US-09-109-886-1	Sequence 1, Appl1	C 818	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 746	12.8	71.1	2271	3	US-08-910-820-8	Sequence 8, Appl1	C 819	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 747	12.8	71.1	2271	3	US-09-844-908-8	Sequence 8, Appl1	C 820	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 748	12.8	71.1	2309	4	US-09-367-895-39	Sequence 39, Appl1	C 821	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 749	12.8	71.1	2478	1	US-08-339-578-1	Sequence 1, Appl1	C 822	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 750	12.8	71.1	2592	4	US-08-481-625-1	Sequence 1, Appl1	C 823	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 751	12.8	71.1	2609	4	US-09-221-017B-642	Sequence 642, Ap	C 824	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 752	12.8	71.1	2609	4	US-09-437-568A-1	Sequence 1, Appl1	C 825	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 753	12.8	71.1	2661	4	US-09-270-767-10462	Sequence 10462, A	C 826	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 754	12.8	71.1	2661	4	US-08-369-822C-20	Sequence 20, Appl1	C 827	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 755	12.8	71.1	2658	3	US-08-582-776C-33	Sequence 33, Appl1	C 828	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 756	12.8	71.1	2658	3	US-08-434-831B-58	Sequence 58, Appl1	C 829	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 757	12.8	71.1	2694	1	US-08-465-995A-1	Sequence 1, Appl1	C 830	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1

831	12.8	71.1	7454	4	US-08-836-687B-29	Sequence 29, Appl	504	12.4	68.9	515	4	US-09-621-976-1563	Sequence 1563, Ap
832	12.8	71.1	8424	4	US-09-543-681A-108	Sequence 808, Appl	505	12.4	68.9	519	4	US-09-621-976-10460	Sequence 10460, A
833	12.8	71.1	8924	3	US-08-369-822C-19	Sequence 19, Appl	506	12.4	68.9	522	4	US-09-621-976-1564	Sequence 1564, A
834	12.8	71.1	8910	3	US-08-779-764A-1	Sequence 1, Appl	507	12.4	68.9	537	4	US-09-248-796A-1162	Sequence 1162, Ap
835	12.8	71.1	8910	3	US-08-583-776C-19	Sequence 19, Appl	508	12.4	68.9	540	3	US-08-714-918-56	Sequence 56, Appl
836	12.8	71.1	8910	3	US-08-434-831B-19	Sequence 19, Appl	509	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
837	12.8	71.1	8910	3	US-09-563-456-1	Sequence 1, Appl	510	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
838	12.8	71.1	10470	4	US-08-956-171E-20	Sequence 20, Appl	511	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
839	12.8	71.1	10470	4	US-08-781-986A-20	Sequence 20, Appl	512	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
840	12.8	71.1	10470	4	US-08-956-171E-20	Sequence 20, Appl	513	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
841	12.8	71.1	11014	4	US-08-956-171E-91	Sequence 91, Appl	514	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
842	12.8	71.1	11580	3	US-09-334-220-4	Sequence 4, Appl	515	12.4	68.9	546	3	US-09-134-001C-55	Sequence 55, Appl
843	12.8	71.1	14113	3	US-09-223-134-1	Sequence 1, Appl	516	12.4	68.9	546	3	US-09-134-001C-55	Sequence 55, Appl
844	12.8	71.1	14113	3	US-08-992-801-1	Sequence 1, Appl	517	12.4	68.9	549	4	US-09-252-991A-3125	Sequence 3125, Ap
845	12.8	71.1	14113	3	US-09-223-535-1	Sequence 1, Appl	518	12.4	68.9	549	4	US-09-252-991A-14225	Sequence 14225, A
846	12.8	71.1	17656	4	US-09-433-579-3	Sequence 3, Appl	519	12.4	68.9	549	4	US-09-248-796A-3140	Sequence 268, Ap
847	12.8	71.1	17656	4	US-08-311-731A-122	Sequence 122, App	520	12.4	68.9	599	4	US-09-621-976-268	Sequence 1, Appl
848	12.8	71.1	74962	4	US-09-685-853A-3	Sequence 3, Appl	521	12.4	68.9	602	1	US-08-229-393-1	Sequence 139, App
849	12.8	71.1	98916	4	US-09-815-055-3	Sequence 3, Appl	522	12.4	68.9	606	4	US-09-489-039A-4415	Sequence 4415, Ap
850	12.8	71.1	129908	4	US-09-585-858-1	Sequence 1, Appl	523	12.4	68.9	618	4	US-09-669-751-139	Sequence 139, App
851	12.8	71.1	169998	4	US-09-676-610B-24	Sequence 24, Appl	524	12.4	68.9	618	4	US-09-248-796A-3806	Sequence 3806, Ap
852	12.8	71.1	197496	4	US-09-877-177A-10	Sequence 10, Appl	525	12.4	68.9	618	4	US-09-110-279-693	Sequence 693, Appl
853	12.8	71.1	269223	4	US-09-596-002-41	Sequence 41, Appl	526	12.4	68.9	627	4	US-09-710-279-1335	Sequence 1335, Ap
854	12.8	71.1	580073	4	US-08-545-552B-1	Sequence 1, Appl	527	12.4	68.9	633	3	US-09-134-001C-578	Sequence 578, App
855	12.8	71.1	640681	4	US-09-790-988-1	Sequence 1, Appl	528	12.4	68.9	633	3	US-09-248-796A-906	Sequence 906, App
856	12.8	71.1	1230025	4	US-08-198-452A-1	Sequence 1, Appl	529	12.4	68.9	643	4	US-09-621-976-578	Sequence 578, App
857	12.4	68.9	20	4	US-09-198-452A-5994	Sequence 5994, Ap	530	12.4	68.9	657	4	US-09-248-796A-1548	Sequence 1548, Ap
858	12.4	68.9	65	3	US-09-191-852-17	Sequence 17, Appl	531	12.4	68.9	658	4	US-09-799-451-697	Sequence 697, App
859	12.4	68.9	65	3	US-08-817-906-17	Sequence 17, Appl	532	12.4	68.9	660	4	US-09-248-796A-4480	Sequence 4480, Ap
860	12.4	68.9	65	5	PCT-US95-13376-17	Sequence 17, Appl	533	12.4	68.9	666	4	US-09-328-352-589	Sequence 589, App
861	12.4	68.9	87	4	US-09-513-999C-13601	Sequence 13601, A	534	12.4	68.9	667	4	US-09-270-767-114063	Sequence 14063, A
862	12.4	68.9	189	4	US-09-248-796A-11625	Sequence 11625, A	535	12.4	68.9	678	3	US-09-232-479-19	Sequence 19, Appl
863	12.4	68.9	220	4	US-09-513-999C-10749	Sequence 10749, A	536	12.4	68.9	678	4	US-09-248-796A-1620	Sequence 1620, Ap
864	12.4	68.9	221	1	US-08-608-789-12	Sequence 12, Appl	537	12.4	68.9	684	4	US-09-453-702B-35	Sequence 35, Appl
865	12.4	68.9	221	2	US-09-111-348-12	Sequence 12, Appl	538	12.4	68.9	686	4	US-09-453-702B-35	Sequence 35, Appl
866	12.4	68.9	225	4	US-09-248-796A-1153	Sequence 1153, Ap	539	12.4	68.9	686	4	US-09-453-702B-35	Sequence 35, Appl
867	12.4	68.9	279	4	US-09-248-796A-8913	Sequence 8913, Ap	540	12.4	68.9	701	4	US-09-641-638-179	Sequence 179, App
868	12.4	68.9	286	4	US-09-313-294A-4525	Sequence 4525, Ap	541	12.4	68.9	701	4	US-10-170-097-179	Sequence 179, App
869	12.4	68.9	286	4	US-09-513-999C-14678	Sequence 14678, A	542	12.4	68.9	705	4	US-09-328-352-860	Sequence 860, App
870	12.4	68.9	291	4	US-09-248-796A-5206	Sequence 5206, Ap	543	12.4	68.9	708	4	US-09-252-991A-3024	Sequence 3024, Ap
871	12.4	68.9	292	4	US-09-313-294A-5755	Sequence 5755, Ap	544	12.4	68.9	711	4	US-09-252-991A-2829	Sequence 2829, Ap
872	12.4	68.9	306	4	US-09-313-294A-5012	Sequence 5012, A	545	12.4	68.9	717	4	US-09-248-796A-4748	Sequence 4748, Ap
873	12.4	68.9	311	4	US-09-513-999C-15612	Sequence 15612, A	546	12.4	68.9	717	4	US-09-107-532A-1165	Sequence 1165, Ap
874	12.4	68.9	312	3	US-09-191-852-20	Sequence 20, Appl	547	12.4	68.9	735	4	US-09-248-796A-878	Sequence 878, App
875	12.4	68.9	312	3	US-08-817-906-20	Sequence 20, Appl	548	12.4	68.9	792	4	US-09-248-796A-878	Sequence 878, App
876	12.4	68.9	312	5	PCT-US95-13376-20	Sequence 20, Appl	549	12.4	68.9	818	4	US-08-956-171E-742	Sequence 742, App
877	12.4	68.9	332	4	US-09-270-767-31505	Sequence 31505, A	550	12.4	68.9	818	4	US-08-781-986A-742	Sequence 742, App
878	12.4	68.9	336	4	US-09-489-039A-6517	Sequence 6517, Ap	551	12.4	68.9	833	4	US-09-553-110-282	Sequence 110, App
879	12.4	68.9	351	3	US-09-134-001C-2503	Sequence 2503, Ap	552	12.4	68.9	882	4	US-09-110-279-2645	Sequence 2645, Ap
880	12.4	68.9	366	4	US-09-248-796A-2185	Sequence 2185, Ap	553	12.4	68.9	882	4	US-09-270-767-1518	Sequence 1518, A
881	12.4	68.9	367	4	US-09-702-705-945	Sequence 945, App	554	12.4	68.9	890	4	US-09-328-352-1956	Sequence 1956, Ap
882	12.4	68.9	367	4	US-09-736-457-945	Sequence 945, App	555	12.4	68.9	897	5	US-09-248-796A-3605	Sequence 3605, Ap
883	12.4	68.9	367	4	US-09-614-124B-945	Sequence 945, App	556	12.4	68.9	900	4	PCT-US96-05320A-403	Sequence 403, App
884	12.4	68.9	367	4	US-09-671-323-945	Sequence 945, App	557	12.4	68.9	900	4	US-09-248-796A-6667	Sequence 6667, Ap
885	12.4	68.9	367	4	US-09-658-824-945	Sequence 945, App	558	12.4	68.9	915	4	US-09-248-796A-710	Sequence 710, App
886	12.4	68.9	375	4	US-09-513-999C-322	Sequence 322, App	559	12.4	68.9	918	4	US-09-248-796A-6594	Sequence 6594, Ap
887	12.4	68.9	380	4	US-09-513-999C-9573	Sequence 9573, Ap	560	12.4	68.9	921	4	US-09-270-767-1519	Sequence 1519, A
888	12.4	68.9	385	4	US-09-270-767-1425	Sequence 1425, Ap	561	12.4	68.9	921	4	US-09-270-767-1611	Sequence 1611, A
889	12.4	68.9	385	4	US-09-270-767-16707	Sequence 16707, A	562	12.4	68.9	945	4	US-09-328-352-1956	Sequence 1956, Ap
890	12.4	68.9	392	4	US-09-621-976-15569	Sequence 15569, A	563	12.4	68.9	945	4	US-09-248-796A-3605	Sequence 3605, Ap
891	12.4	68.9	420	4	US-09-248-796A-5722	Sequence 5722, Ap	564	12.4	68.9	975	4	US-09-252-991A-5535	Sequence 5535, Ap
892	12.4	68.9	435	3	US-09-134-001C-123	Sequence 123, App	565	12.4	68.9	1001	4	US-09-641-638-180	Sequence 180, App
893	12.4	68.9	440	4	US-09-513-999C-13884	Sequence 13884, A	566	12.4	68.9	1001	4	US-10-170-097-180	Sequence 180, App
894	12.4	68.9	441	4	US-09-252-991A-13884	Sequence 34591, A	567	12.4	68.9	1005	4	US-09-248-796A-3263	Sequence 3263, Ap
895	12.4	68.9	459	4	US-09-270-767-7119	Sequence 7119, Ap	568	12.4	68.9	1026	4	US-09-107-532A-3333	Sequence 3333, Ap
896	12.4	68.9	459	4	US-09-270-767-22401	Sequence 22401, A	569	12.4	68.9	1032	4	US-08-956-171E-247	Sequence 247, App
897	12.4	68.9	462	4	US-09-513-999C-22848	Sequence 22848, A	570	12.4	68.9	1032	4	US-08-781-986A-247	Sequence 247, App
898	12.4	68.9	468	4	US-09-248-796A-3653	Sequence 3653, Ap	571	12.4	68.9	1066	4	US-09-107-532A-163	Sequence 163, App
899	12.4	68.9	474	4	US-09-621-976-1566	Sequence 1566, Ap	572	12.4	68.9	1074	4	US-09-270-767-2557	Sequence 2557, Ap
900	12.4	68.9	477	4	US-09-248-796A-5071	Sequence 5071, Ap	573	12.4	68.9	1090	4	US-09-270-767-15180	Sequence 15180, A
901	12.4	68.9	478	4	US-09-621-976-10037	Sequence 10037, A	574	12.4	68.9	1101	4	US-09-248-796A-863	Sequence 863, App
902	12.4	68.9	483	4	US-09-583-110-2569	Sequence 2569, Ap	575	12.4	68.9	1107	4	US-08-956-171E-638	Sequence 638, App
903	12.4	68.9	507	4	US-09-248-796A-7370	Sequence 7370, Ap	576	12.4	68.9	1110	4	US-08-781-986A-658	Sequence 658, App
												US-09-248-796A-1340	Sequence 1340, Ap

977 12.4 68.9 1155 4 US-09-252-991A-13685 Sequence 13685, A
C 978 12.4 68.9 1158 4 US-09-248-796A-2478 Sequence 2478, Ap
979 12.4 68.9 1164 3 US-09-134-001C-1461 Sequence 1461, Ap
C 980 12.4 68.9 1170 4 US-09-248-796A-3621 Sequence 3621, Ap
C 981 12.4 68.9 1182 4 US-09-248-796A-4758 Sequence 4758, Ap
C 982 12.4 68.9 1183 4 US-09-684-385-9 Sequence 9, Appl1
983 12.4 68.9 1203 4 US-09-248-796A-3748 Sequence 3748, Ap
984 12.4 68.9 1205 2 US-08-680-385-7 Sequence 7, Appl1
985 12.4 68.9 1214 4 US-09-506-286B-15 Sequence 15, Appl1
986 12.4 68.9 1214 4 US-09-506-286B-18 Sequence 18, Appl1
987 12.4 68.9 1214 4 US-09-762-861B-15 Sequence 15, Appl1
988 12.4 68.9 1214 4 US-09-762-861B-18 Sequence 18, Appl1
989 12.4 68.9 1214 4 US-10-065-133A-15 Sequence 15, Appl1
990 12.4 68.9 1214 4 US-10-065-133A-18 Sequence 18, Appl1
991 12.4 68.9 1215 4 US-09-328-352-3884 Sequence 3884, Ap
992 12.4 68.9 1229 1 US-08-180-209B-54 Sequence 54, Appl1
993 12.4 68.9 1229 3 US-08-474-853-54 Sequence 54, Appl1
994 12.4 68.9 1229 3 US-09-166-205B-54 Sequence 54, Appl1
995 12.4 68.9 1229 5 PCT-US94-02629-54 Sequence 54, Appl1
996 12.4 68.9 1241 4 US-09-506-286B-13 Sequence 13, Appl1
997 12.4 68.9 1241 4 US-09-506-286B-16 Sequence 16, Appl1
998 12.4 68.9 1241 4 US-09-762-861B-13 Sequence 13, Appl1
999 12.4 68.9 1241 4 US-09-762-861B-16 Sequence 16, Appl1
1000 12.4 68.9 1241 4 US-10-065-133A-13 Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-134-000C-2523
; Sequence 2523, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055, 778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2523
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2523

Query Match 85.6%; Score 15.4; DB 4; Length 534;
Best Local Similarity 94.1%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTCGC 18
DB 465 AGACACACAGATTCGC 481

RESULT 2
US-08-968-563-7
; Sequence 7, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY W. WHITED
; APPLICANT: VASANTHA NAGARAJAN

APPLICANT: MARK S. PAYNE
APPLICANT: STEPHEN K. PICCAGGIO
APPLICANT: RAMESCH V. NAIR
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898

ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
STREET: 1870 SOUTH WINTON ROAD
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT2
US-08-968-563-7

Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTCGC 18
DB 2154 AGACACACAGATTCGC 2170

RESULT 3
US-08-969-683A-7
; Sequence 7, Application US/08969683A
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; FILE REFERENCE: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
STREET: 1870 South Winton Road
CITY: Rochester
STATE: NY

COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,683A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/20873
FILING DATE: 13-NOV-1997
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gialster, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT2
US-08-969-683A-7

Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGGC 18
DB 2154 AGAACAACAAGATTGGC 2170

RESULT 4
US-09-297-928-3
Sequence 3, Application US/09297928
Patent No. 6358716
GENERAL INFORMATION:
APPLICANT: BULTHUIS, BEN A.
GATENBY, ANTHONY A.
HAYNIE, SHARON L.
HSU, AMY K.
LARBAU, RICHARD D.
TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
GLYCEROL BY RECOMBINANT
ORGANISMS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/297,928

FILING DATE: 11-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/03602
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9981-P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
TELEX: 6717325
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-297-928-3

Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGGC 18
DB 2154 AGAACAACAAGATTGGC 2170

RESULT 5
US-09-276-599-4/c
Sequence 4, Application US/09276599
Patent No. 6380459
GENERAL INFORMATION:
APPLICANT: Perera, J. Ranjan
TITLE OF INVENTION: Composition and methods for the
production of gene expression
FILE REFERENCE: 11000,1036
CURRENT APPLICATION NUMBER: US/09/276,599
NUMBER OF SEQ ID NOS: 21
CURRENT FILING DATE: 1999-03-25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 485
TYPE: DNA
ORGANISM: Pinus radiata
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(431)
NAME/KEY: TATA signal
LOCATION: (350)...(356)
NAME/KEY: CAAT signal
LOCATION: (326)...(333)
US-09-276-599-4

Query Match 83.3%; Score 15; DB 3; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACACAAGATTGGC 18
DB 409 AACACAAGATTGGC 395

RESULT 6
US-09-598-401c-4/c
Sequence 4, Application US/09598401C
Patent No. 6596925
GENERAL INFORMATION:

```
/ APPLICANT: Ferreira, J. Ranjan
/ APPLICANT: Sagleton, Clare
/ APPLICANT: Rice, Stephen J.
/ TITLE OF INVENTION: Compositions and Methods for the
/ TITLE OF INVENTION: Modification of Gene Expression
/ FILE REFERENCE: 11000.10362
/ CURRENT APPLICATION NUMBER: US/09/598,401C
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
/ PRIOR FILING DATE: 1999-07-30
/ PRIOR APPLICATION NUMBER: PCT/N200/00018
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
/ PRIOR FILING DATE: 1999-03-25
/ NUMBER OF SEQ ID NOS: 120
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 485
/ TYPE: DNA
/ ORGANISM: Pinus radiata
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: (1)...(431)
/ FEATURE:
/ NAME/KEY: TATA signal
/ LOCATION: (350)...(356)
/ FEATURE:
/ NAME/KEY: CAT signal
/ LOCATION: (326)...(333)
/ US-09-598-401C-4
```

Query Match 83.3%; Score 15; DB 4; Length 485;

Best Local Similarity 100.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AACACACAGATTGCG 18

Db 409 AACACACAGATTGCG 395

```
RESULT 7
/ US-09-248-796A-6542
/ Sequence 6542, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 6542
/ LENGTH: 1359
/ TYPE: DNA
/ ORGANISM: Candida albicans
/ US-09-248-796A-6542
```

Query Match 83.3%; Score 15; DB 4; Length 1359;

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGACACACAGATTG 16

Db 696 AGACACACAGATTG 710

RESULT 8
US-09-495-050A-15

```
/ Sequence 15, Application US/09495050A
/ Patent No. 6492505
/ GENERAL INFORMATION:
/ APPLICANT: Roopa, Reddy
/ APPLICANT: Giegler, Karl, J.
/ APPLICANT: Au-Yang, Janice
/ TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED PR
/ FILE REFERENCE: PA-0013 US
/ CURRENT APPLICATION NUMBER: US/09/495,050A
/ CURRENT FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/118,318
/ PRIOR FILING DATE: February 1, 1999
/ NUMBER OF SEQ ID NOS: 305
/ SOFTWARE: PERL Program
/ SEQ ID NO 15
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6492505 227799CBI
/ US-09-495-050A-15
```

Query Match 82.2%; Score 14.8; DB 4; Length 676;

Best Local Similarity 88.9%; Pred. No. 1.9e+02; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGACACACAGATTGCG 18

Db 30 GAGACACACAGATTGCG 47

```
RESULT 9
/ US-09-270-767-10453
/ Sequence 10453, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 10453
/ LENGTH: 1098
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
/ US-09-270-767-10453
```

Query Match 82.2%; Score 14.8; DB 4; Length 1098;

Best Local Similarity 88.9%; Pred. No. 2.1e+02; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGACACACAGATTGCG 18

Db 660 GTGACACACAGATTGCG 677

```
RESULT 10
/ US-09-328-352-3042/C
/ Sequence 3042, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 3042
/ LENGTH: 2727
```

TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3042

Query Match 82.2%; Score 14.8; DB 4; Length 2727;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 2114 GAGAACACAGATTGCG 2097

RESULT 11

US-08-804-227C-7
Sequence 7, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Kuhnstock, Stuart A.

APPLICANT: Rostock, Paul R., Jr.

APPLICANT: Sutton, Kimberly L.

TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: THOMAS G. PLANT 1501

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII(DOS) Text only

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804.227C

FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.

REGISTRATION NUMBER: 35,784

REFERENCE/DOCKET NUMBER: X-8231

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 4437 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 350..14002

FEATURE:

NAME/KEY: CDS

LOCATION: 14046..20036

FEATURE:

NAME/KEY: CDS

LOCATION: 20110..31284

Qy 1 GAGAACACAGATTGCG 18
Db 14055 GAGAACACAGATTGCG 14072

RESULT 12
US-08-804-198-1
Sequence 1, Application US/08804198
Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Bargett, Stanley G.

APPLICANT: Kuhnstock, Stuart A.

APPLICANT: Rao, Nagaraja R.

APPLICANT: Richardson, Mark A.

APPLICANT: Rostock, Paul R., Jr.

TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: PAUL R. CANTRELL 1138

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,198

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.

REGISTRATION NUMBER: 36,470

REFERENCE/DOCKET NUMBER: P9113

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4437 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 350..14002

FEATURE:

NAME/KEY: CDS

LOCATION: 14046..20036

FEATURE:

NAME/KEY: CDS

LOCATION: 20110..31284

FEATURE:

NAME/KEY: CDS

LOCATION: 36155..41830

US-08-804-198-1

Query Match 82.2%; Score 14.8; DB 2; Length 44377;
Best Local Similarity 88.9%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGCG 18
Db 14055 GAGAACACAGATTGCG 14072

Query Match 82.2%; Score 14.8; DB 2; Length 44377;
Best Local Similarity 88.9%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      82.2%; Score 14.8; DB 3; Length 4403765;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGAACACACAGATTCCG 18
DB      3556350 GAGAACACAAAGTTCCG 3556367

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      82.2%; Score 14.8; DB 3; Length 4411529;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGAACACACAGATTCCG 18
DB      3561497 GAGAACACAAAGTTCCG 3561514

RESULT 15
US-09-248-796A-13022/C
; Sequence 13022, Application US/09248796A
; Patent No. 6747117
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13022
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-13022

Query Match      80.0%; Score 14.4; DB 4; Length 204;
Best Local Similarity 93.8%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 17
DB      181 AGAACACACATTTCCG 166

RESULT 16
US-09-252-991A-2932
; Sequence 2932, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2932
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2932

Query Match      80.0%; Score 14.4; DB 4; Length 591;
Best Local Similarity 93.8%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 GAGAACACAGATTCCG 18
DB      11 GAGAACACAGATTCCC 26

RESULT 17
US-09-252-991A-2843
; Sequence 2843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2843
```

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rn1

Page 13

LENGTH: 684
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2843

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 684;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACACACAGATTCCG 18
DB 186 GAACACACAGATTCCG 201

RESULT 18
US-09-270-767-13143/C
Sequence 13143, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13143
LENGTH: 771
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-13143

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 771;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 17
DB 496 AGAACACACAGATTCCG 481

RESULT 19
US-09-577-934A-1
Sequence 1, Application US/09577934A
Patent No. 6590087
GENERAL INFORMATION:
APPLICANT: Bishai, William
TITLE OF INVENTION: WMD, AN ESSENTIAL CELL DIVISION GENE FROM MYCOBACTERIA
FILE REFERENCE: Bishai
CURRENT APPLICATION NUMBER: US/09/577.934A
CURRENT FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 783
TYPE: DNA
ORGANISM: Mycobacterium smegmatis
FEATURE:
NAME/KEY: CDS
LOCATION: (188)..(577)
US-09-577-934A-1

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 783;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACACACAGATTCCG 18
DB 753 GAACACACAGATTCCG 768

RESULT 20

US-09-252-991A-3123/C
Sequence 3123, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074.788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094.190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3123
LENGTH: 867
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3123

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 867;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACACACAGATTCCG 18
DB 637 GAACACACAGATTCCG 622

RESULT 21
US-08-956-171E-771/C
Sequence 771, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009.861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781.986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION/DOCKET NUMBER: 46,789
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 771:
SEQUENCE CHARACTERISTICS:

LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 771:
US-08-956-171E-771

Query Match 80.0%; Score 14.4; DB 4; Length 1462;
Best Local Similarity 93.8%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTC 17
DB 816 AGAACACACAGATTTC 801

RESULT 22
US-08-781-986A-771/C
Sequence 771, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:
APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 771:
SEQUENCE CHARACTERISTICS:
LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-771

Query Match 80.0%; Score 14.4; DB 4; Length 1462;
Best Local Similarity 93.8%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTC 17
DB 816 AGAACACACAGATTTC 801

RESULT 23
US-09-620-312D-516
Sequence 516, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aldong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dantui
APPLICANT: Wang, Zhwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: 784C1P2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 516
LENGTH: 1823
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (114)..(800)
US-09-620-312D-516

Query Match 80.0%; Score 14.4; DB 4; Length 1823;
Best Local Similarity 93.8%; Pred. No. 3.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTC 17
DB 547 AGAACACACAGATTTC 562

RESULT 24
US-09-556-601-25
Sequence 25, Application US/09556601
Patent No. 6381599
GENERAL INFORMATION:
APPLICANT: Jefferies, Thomas W.
TITLE OF INVENTION: SHAW-SENSITIVE TERMINAL OXIDASE GENE FROM
FILE REFERENCE: 96429.9074
CURRENT APPLICATION NUMBER: US/09/556,601
CURRENT FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Word 97 (DOS text format)
SEQ ID NO 25
LENGTH: 2853
TYPE: DNA
ORGANISM: Pichia stipitis
FEATURE:
NAME/KEY: CDS
LOCATION: (954)..(2027)
US-09-556-601-25

Query Match 80.0%; Score 14.4; DB 3; Length 2853;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTC 16

Db 1767 GAGAACATCAAGATTTC 1782

RESULT 25
US-09-907-794A-310

; Sequence 310, Application US/09907794A

; Patent No. 6635468

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Kjaer, Jennie P.

; APPLICANT: Maier, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-310
Query Match 80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGAACATCAAGATTTC 17

Db 2017 AGAACATCAAGATTTC 2032

RESULT 26

US-09-905-125A-310

; Sequence 310, Application US/09905125A

; Patent No. 6664376

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Kjaer, Jennie P.

; APPLICANT: Maier, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

; APPLICANT: Wood, William, I.

```

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-09-905-125A-310

Query Match      80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AGACACACAGATTG 17
Db      2017 AGACACACAGATTG 2032

RESULT 27
US-09-902-775A-310
; Sequence 310, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Grimsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-09-902-775A-310

Query Match      80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AGACACACAGATTG 17
Db      2017 AGACACACAGATTG 2032

RESULT 28
US-09-906-700-310
; Sequence 310, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Grimsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
```


CURRENT APPLICATION NUMBER: US/09/906,700
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 310
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-09-906-700-310

Query Match 80.0%; Score 14.4; DB 4;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACAGATTG 17
Db 2017 AGAACACAGATTG 2032

RESULT 29
US-10-140-002-369
Sequence 369, Application US/10140002
Patent No. 6725730
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William

APPLICANT: Zhang, Zenit
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 369
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-369

Query Match 80.0%; Score 14.4; DB 4;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACAGATTG 17
Db 2017 AGAACACAGATTG 2032

RESULT 30
US-09-903-603A-310
Sequence 310, Application US/09903603A
Patent No. 676795
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: GNE 1618P2C12
CURRENT APPLICATION NUMBER: US/09/903,603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05 PCT/US99/28214
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29 PCT/US99/28313
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30 PCT/US99/28564
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02 PCT/US99/28565
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02 PCT/US99/30095
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16 PCT/US99/30911
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20 PCT/US99/30999
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20 PCT/US00/00219
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO: 310
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-09-903-603A-310

Query Match 80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTG 17
DB 2017 AGAACACACAGATTAG 2032

RESULT 31
US-08-936-171E-228
Sequence 228, Application US/08956171E
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunesh
Gail H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 228:
SEQUENCE CHARACTERISTICS:
LENGTH: 4488 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 228:
US-08-956-171E-228

Query Match 80.0%; Score 14.4; DB 4; Length 4488;
Best Local Similarity 93.8%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTG 16
DB 111 GAGAACACACAGATTG 126

RESULT 32
US-08-781-986A-228
Sequence 228, Application US/08781986A
Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunesh
Gail H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248BP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 228:

SEQUENCE CHARACTERISTICS:
LENGTH: 4488 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-228

Query Match 80.0%; Score 14.4; DB 4; Length 4488;
Best Local Similarity 93.8%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTG 16
DB 111 GAGAACACACAGATTG 126

RESULT 33

```
US-09-345-882-1
Sequence 1, Application US/09345882
Patent No. 6399373
GENERAL INFORMATION:
APPLICANT: Bouquelere, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FILE REFERENCE: GENSET.031A
CURRENT APPLICATION NUMBER: US/09/345.882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
```

```
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match      80.0%; Score 14.4; DB 3; Length 162450;
Best Local Similarity 93.8%; Pred. No. 5.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGAACAAACAAGATTC 16
Db      95472 GACAAACAAGATTC 95487

RESULT 34
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
```

```
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "t" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      80.0%; Score 14.4; DB 3; Length 4403765;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGAACAAACAAGATTC 16
Db      2998437 GCGAACAAACAAGATTC 2998422

RESULT 35
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      80.0%; Score 14.4; DB 3; Length 4411529;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGAACAAACAAGATTC 16
Db      3003531 GCGAACAAACAAGATTC 3003516

RESULT 36
US-09-252-991A-15780
Sequence 15780, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUSINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 15780
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15780

Query Match 77.8%; Score 14; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGAT 14
|||
DB 11 GAGACACACAGAT 24

RESULT 37
US-09-248-796A-662
; Sequence 662, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO: 662
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-662

Query Match 77.8%; Score 14; DB 4; Length 792;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATT 15
|||
DB 777 AGAACACACAGATT 790

RESULT 38
US-09-252-991A-15910/C
; Sequence 15910, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 15910
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15910

Query Match 77.8%; Score 14; DB 4; Length 825;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGAT 14
|||
DB 732 GAGACACACAGAT 719

RESULT 39
US-09-248-796A-10471
; Sequence 10471, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO: 10471
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-10471

Query Match 77.8%; Score 14; DB 4; Length 969;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATT 15
|||
DB 252 AGAACACACAGATT 265

RESULT 40
US-09-252-991A-15881/C
; Sequence 15881, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 15881
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15881

Query Match 77.8%; Score 14; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGAT 14
|||
DB 1018 GAGACACACAGAT 1005

RESULT 41
US-09-252-991A-15836
; Sequence 15836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15836
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-15836

Query Match      77.8%; Score 14; DB 4; Length 1317;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GAGAACCAACAAGAT 14
Db      1293 GAGAACCAACAAGAT 1306

RESULT 42
US-09-252-991A-15811
; Sequence 15811, Application US/09252991A
; Patent No. 6551785
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15811
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-15811

Query Match      77.8%; Score 14; DB 4; Length 2058;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GAGAACCAACAAGAT 14
Db      19 GAGAACCAACAAGAT 32

RESULT 43
US-09-809-665A-39
; Sequence 39, Application US/09809665A
; Patent No. 6790950
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28141/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
```

```

; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2628
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (326)..(766)
; OTHER INFORMATION: mloc
; US-09-809-665A-39

Query Match      77.8%; Score 14; DB 4; Length 2628;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 AGAACCAACAAGAT 15
Db      406 AGAACCAACAAGAT 419

RESULT 44
US-08-956-171E-1643
; Sequence 1643, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Farnon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1643:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1643:
US-08-956-171E-1643

Query Match      76.7%; Score 13.8; DB 4; Length 66;
Best Local Similarity 88.2%; Pred. No. 4.7e+02;
```

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACAAAGATTCCG 18
 Db 42 AGAACATATAGATCCG 58

RESULT 45

US-08-781-986A-1643
 ; Sequence 1643, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunach
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 1643:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 66 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-781-986A-1643

Query Match 76.7%; Score 13.8; DB 4; Length 66;

Best Local Similarity 88.2%; Pred. No. 4.7e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACAAAGATTCCG 18
 Db 42 AGAACATATAGATCCG 58

RESULT 46

US-09-222-575-44
 ; Sequence 44, Application US/09222575
 ; Patent No. 6387697
 ; GENERAL INFORMATION:
 ; APPLICANT: Yugu, Jiang
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jianshun
 ; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
 ; TITLE OF INVENTION: and Methods for Their Use
 ; FILE REFERENCE: 210121.470
 ; CURRENT APPLICATION NUMBER: US/09/222,575
 ; CURRENT FILING DATE: 1998-12-28
 ; NUMBER OF SEQ ID NOS: 174

SOFTWARE: FastSeq for Windows Version 3.0

QY 2 AGAACACAAAGATTCCG 18
 Db 45 AGAACACAAAGATTCCG 61

Query Match

Best Local Similarity 88.2%; Pred. No. 5.2e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACAAAGATTCCG 18
 Db 45 AGAACACAAAGATTCCG 61

RESULT 47

US-09-389-681-44
 ; Sequence 44, Application US/09389681A
 ; Patent No. 6518237
 ; GENERAL INFORMATION:
 ; APPLICANT: Yugu, Jiang
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jianshun
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.470C3
 ; CURRENT APPLICATION NUMBER: US/09/389,681A
 ; CURRENT FILING DATE: 1999-09-02
 ; NUMBER OF SEQ ID NOS: 463
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 160
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-389-681-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;
 Best Local Similarity 88.2%; Pred. No. 5.2e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACAAAGATTCCG 18
 Db 45 AGAACACAAAGATTCCG 61

RESULT 48

US-09-620-405B-44
 ; Sequence 44, Application US/09620405B
 ; Patent No. 6528054
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Hepler, William T.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.470C8
 ; CURRENT APPLICATION NUMBER: US/09/620,405B
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 495
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 160
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-620-405B-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;
 Best Local Similarity 88.2%; Pred. No. 5.2e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 45 AGAACACACAGATCCAC 61

RESULT 49

US-09-339-338-44
; Sequence 44, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-339-338-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;

Best Local Similarity 88.2%; Pred. No. 5.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 45 AGAACACACAGATCCAC 61

RESULT 50

US-09-433-826B-44
; Sequence 44, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;

Best Local Similarity 88.2%; Pred. No. 5.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
Db 45 AGAACACACAGATCCAC 61

Search completed: December 3, 2004, 05:51:27
Job time : 123.684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 01:41:45 ; Search time 2282.45 Seconds
(without alignments)
287.374 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18

Sequence: 1 gagagacacaaagattcgc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 segs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hlc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	94.4	164	9	CG665075 OST452877
2	17	94.4	383	9	CG559318 OST178045
3	17	94.4	936	9	CNS07414 CNS07414
4	16.4	91.1	179	4	BG410386 947045H08
5	16.4	91.1	357	6	CD460790 FG09_01n0
6	16.4	91.1	481	8	AO210247 HS_31.5 B
7	16.4	91.1	504	6	CD557825 EtEST6E59
8	16.4	91.1	514	4	BM076070 MEST364-D
9	16.4	91.1	768	7	CO817404 PA_SEA001
10	16.4	91.1	1138	4	B1823206 603039469
11	16	88.9	466	2	BE574801 H19_TripH
12	16	88.9	426	9	CE085246 LiGR-gss-
13	16	88.9	646	7	CK130476 AT08181.3
14	16	88.9	690	8	BZ178809 CH230.443
15	16	88.9	744	8	CC082521 CSU-K31T.
16	16	88.9	822	5	BU017470 QHE15P17.
17	15.4	85.6	229	5	BU018252 QHE18B18.
18	15.4	85.6	229	5	BU018252 QHE18B18.
19	15.4	85.6	269	5	BU016050 QHE11P07.
20	15.4	85.6	270	6	CD944524 RDJ_72 Ge
21	15.4	85.6	270	6	CD946927 REV_27 Ge
22	15.4	85.6	270	6	CD967114 SES_132 G
23	15.4	85.6	270	6	CD967248 SES_61 Ge
24	15.4	85.6	273	2	AW977559 EST1389668

25	15.4	85.6	284	5	BO910769 CHA15C10.
26	15.4	85.6	285	5	BU016309 QHE12117.
27	15.4	85.6	286	5	BU021781 QHE4E07_Y
28	15.4	85.6	300	6	CA485884 WHE4324_C
29	15.4	85.6	304	2	AW646697 PWE2-SN001
30	15.4	85.6	312	1	AV216565 AV216565
31	15.4	85.6	340	1	AW787371 n2508a1.x
32	15.4	85.6	351	2	AW447131 88282 NAR
33	15.4	85.6	365	1	AU257195 AU257195
34	15.4	85.6	366	1	BZ551766 OGAMP82NC
35	15.4	85.6	367	2	BF851239 IL5-EN008
36	15.4	85.6	380	3	AY440562 Arm1-genes
37	15.4	85.6	383	8	AQ851814 LMAJFV1_1
38	15.4	85.6	400	1	AA994392 ou10g03.s
39	15.4	85.6	402	9	CG722358 CG722358
40	15.4	85.6	406	1	AI280419 q195d08.x
41	15.4	85.6	409	1	AI220277 q573d03.x
42	15.4	85.6	413	1	AA907628 om10F08.s
43	15.4	85.6	419	4	BM173099 900453_AY
44	15.4	85.6	428	1	AA812092 cb40a06.s
45	15.4	85.6	430	8	BZ769778 SALK_1427
46	15.4	85.6	431	5	BP645308 BP645308
47	15.4	85.6	433	4	BG543327 E0832_CH1
48	15.4	85.6	437	4	BM360109 CA_Fa002
49	15.4	85.6	438	1	AA610098 BI0801478
50	15.4	85.6	438	4	BM01478 BI0801478
51	15.4	85.6	466	1	AA772688 AI35B05.s
52	15.4	85.6	474	1	AI266637 qm45b11.x
53	15.4	85.6	478	7	CN261424 170005322
54	15.4	85.6	478	8	BH849056 SALK_0582
55	15.4	85.6	490	4	BU759781 BU759781
56	15.4	85.6	498	9	BH179757 O15_D23-
57	15.4	85.6	498	8	CNS07414 CNS07414
58	15.4	85.6	505	4	BI108381 BI108381
59	15.4	85.6	512	4	BG103963 RH122_38
60	15.4	85.6	516	4	BU760375 BU760375
61	15.4	85.6	517	4	BG103963 RH122_18
62	15.4	85.6	519	8	AQ889206 HS_5539_A
63	15.4	85.6	525	7	CF931508 RTDR3_9_E
64	15.4	85.6	537	2	AW301203 xss57905.s
65	15.4	85.6	537	4	BU765562 BU765562
66	15.4	85.6	539	4	BU753442 BU753442
67	15.4	85.6	543	4	BU772991 BU772991
68	15.4	85.6	548	4	BU126386 BU126386
69	15.4	85.6	551	4	BG355639 EMI_18 DO
70	15.4	85.6	552	4	BU112454 BU112454
71	15.4	85.6	553	6	CD991044 OA23610.Y
72	15.4	85.6	555	1	AI202762 q145d06.x
73	15.4	85.6	561	5	BO664424 HVO3B04u
74	15.4	85.6	565	4	BU112907 BU112907
75	15.4	85.6	565	6	CB015597 HC_d11.42
76	15.4	85.6	566	4	BG465737 RH122_48-
77	15.4	85.6	567	1	AI871351 w181B02.X
78	15.4	85.6	569	1	AI947649 603026A07
79	15.4	85.6	572	5	BO664177 HVO2D11u
80	15.4	85.6	576	1	AI220156 G979B11.x
81	15.4	85.6	579	4	BI508873 BB170002B
82	15.4	85.6	580	4	BM323691 PRC1_22 B
83	15.4	85.6	585	2	AW015551 UI-H-B10P
84	15.4	85.6	587	1	AA775336 A13804.s
85	15.4	85.6	587	7	CO790032 NT008C_DO
86	15.4	85.6	588	6	AW165994 x438G04.x
87	15.4	85.6	588	6	CF022808 CBO19G09.x
88	15.4	85.6	589	4	CE702297 LiGR-gss-
89	15.4	85.6	590	2	BU117751 BU117751
90	15.4	85.6	600	2	AW146888 E14089C12
91	15.4	85.6	603	3	BU477134 BU477134
92	15.4	85.6	622	3	AY613879 Mesenchyt
93	15.4	85.6	624	4	BU753364 BU753364
94	15.4	85.6	625	5	BU573976 PA_Ba000
95	15.4	85.6	625	8	BH745756 G271d08.b
96	15.4	85.6	628	4	BU108119 BU108119
97	15.4	85.6	630	5	BP117787 BP117787

98	15.4	85.6	631	4	BG718181	602696145	171	15	83.3	521	1	AV671114
99	15.4	85.6	657	7	CK511600	rswwd0.00	172	15	83.3	521	2	BE601652
100	15.4	85.6	659	4	BJ770602	Bj770602	173	15	83.3	523	6	CD309974
101	15.4	85.6	663	6	CB332921	py70f07.y	174	15	83.3	539	1	AV55802
102	15.4	85.6	664	7	BQ397128	QGG14M24.	175	15	83.3	545	6	CD823327
103	15.4	85.6	664	7	CP391534	RTDR3_9_E	176	15	83.3	546	5	B0091272
104	15.4	85.6	674	1	A1356774	qy22h06.x	177	15	83.3	550	6	CB261944
105	15.4	85.6	679	8	BH332217	od122d04.	178	15	83.3	553	5	BM453301
106	15.4	85.6	688	4	BG495794	602539050	179	15	83.3	561	5	BM413325
107	15.4	85.6	717	9	AG456384	Mus muscu	180	15	83.3	562	4	BI397065
108	15.4	85.6	737	8	BZ651772	OGANP82TM	181	15	83.3	573	1	AA558907
109	15.4	85.6	739	5	BUT25701	SUNCIH06	182	15	83.3	576	1	BM494170
110	15.4	85.6	746	9	CL667846	PR10156a_	183	15	83.3	593	1	AV561026
111	15.4	85.6	754	9	AG309135	Mus muscu	184	15	83.3	593	1	BM114699
112	15.4	85.6	762	6	CB954477	AGENCOURT	185	15	83.3	594	1	AJ803959
113	15.4	85.6	764	9	CC704364	OGWE174TH	186	15	83.3	599	5	BM429329
114	15.4	85.6	773	7	CC0484316	G00205.B3	187	15	83.3	601	1	AV888715
115	15.4	85.6	775	3	AY223366	Schistoso	188	15	83.3	601	6	CA955346
116	15.4	85.6	777	8	BZ479818	BONCD5TR	189	15	83.3	602	6	CB371276
117	15.4	85.6	800	9	CL457009	ZMMBBD048	190	15	83.3	602	8	AZ995079
118	15.4	85.6	812	7	CK769679	957749.MA	191	15	83.3	603	2	BE039065
119	15.4	85.6	824	7	BZ995959	PURJC70TD	192	15	83.3	609	5	BM473143
120	15.4	85.6	849	9	CC695247	OGVET81TV	193	15	83.3	609	8	BZ633986
121	15.4	85.6	858	8	BZ995953	PURJC70TB	194	15	83.3	612	9	BX185625
122	15.4	85.6	871	9	CR266303	Forward.6	195	15	83.3	615	8	BM465633
123	15.4	85.6	912	9	CL299860	ZMMBBD008	196	15	83.3	618	8	BZ679984
124	15.4	85.6	917	9	CL445921	ZMMBBD046	197	15	83.3	620	6	CA190996
125	15.4	85.6	925	9	CL464318	SA1L_1219	198	15	83.3	622	2	BE214881
126	15.4	85.6	949	2	BF308719	601890011	199	15	83.3	632	5	BQ859935
127	15.4	85.6	994	4	CM042682	OGVCN75TV	200	15	83.3	632	5	BU92061
128	15.4	85.6	1016	4	BM048812	603620189	201	15	83.3	643	1	AV992458
129	15.4	85.6	1033	9	CNS06Y50	T3_end_of	202	15	83.3	650	6	CB883667
130	15.4	85.6	1062	9	CNS02UTS	AL1214921	203	15	83.3	654	5	BD599638
131	15.4	85.6	1141	7	CN970619	18116_123	204	15	83.3	658	6	CA090403
132	15.4	85.6	1292	2	BE967949	60287-830	205	15	83.3	674	5	BM421513
133	15.4	85.6	1312	9	CG749006	P043-1-DO	206	15	83.3	681	5	BM419828
134	15.4	85.6	1461	2	BR971547	602239772	207	15	83.3	682	8	BM425288
135	15.2	84.4	149	1	AX535231	AX535231	208	15	83.3	685	8	BZ636291
136	15.4	83.3	149	1	BP636527	BP636527	209	15	83.3	685	5	CL167604
137	15.4	83.3	273	5	AV880807	AV880807	210	15	83.3	688	5	BM020278
138	15.4	83.3	289	5	BP636527	BP636527	211	15	83.3	692	7	CA090450
139	15.4	83.3	337	6	CD479381	eca01_21m	212	15	83.3	696	8	BM935232
140	15.4	83.3	351	5	BY302652	BY302652	213	15	83.3	698	8	BM940277
141	15.4	83.3	360	5	BP629145	BP629145	214	15	83.3	706	6	CD867400
142	15.4	83.3	367	5	BM072723	BM072723	215	15	83.3	708	6	CA140075
143	15.4	83.3	367	5	BM074459	BM074459	216	15	83.3	717	8	BZ408702
144	15.4	83.3	383	9	CC871300	NDL_121P1	217	15	83.3	721	6	CE235788
145	15.4	83.3	388	9	CL281731	ZMMBBD062	218	15	83.3	721	7	CR285094
146	15.4	83.3	390	9	CG904708	ZMMBBD051	219	15	83.3	723	9	CC874884
147	15.4	83.3	392	8	B35497	HS-1029-A1-	220	15	83.3	727	1	AJ803466
148	15.4	83.3	393	5	BP633218	BP633218	221	15	83.3	730	8	BZ799822
149	15.4	83.3	396	1	AV7884347	AV7884347	222	15	83.3	736	5	BU001781
150	15.4	83.3	407	5	BP587928	BP587928	223	15	83.3	743	9	CL745885
151	15.4	83.3	413	5	BP633053	BP633053	224	15	83.3	747	6	CA266752
152	15.4	83.3	423	5	BM208153	BM208153	225	15	83.3	747	8	BZ467270
153	15.4	83.3	423	5	BM209971	BM209971	226	15	83.3	751	7	CO079632
154	15.4	83.3	425	1	AV819763	AV819763	227	15	83.3	753	5	BM117438
155	15.4	83.3	429	5	BP634162	BP634162	228	15	83.3	759	7	CK120352
156	15.4	83.3	439	8	BH719447	BOMKX04TR	229	15	83.3	760	8	BM455302
157	15.4	83.3	439	8	BM823447	BACBP17-H	230	15	83.3	761	7	CNI80704
158	15.4	83.3	445	5	BP635711	BP635711	231	15	83.3	761	9	CL167603
159	15.4	83.3	448	5	BP593508	BP593508	232	15	83.3	762	9	AC468993
160	15.4	83.3	451	5	BP586034	BP586034	233	15	83.3	763	6	CA166813
161	15.4	83.3	474	1	AV879730	AV879730	234	15	83.3	764	9	BX163759
162	15.4	83.3	477	5	BY256777	BY256777	235	15	83.3	765	6	CA074505
163	15.4	83.3	480	5	AQ721815	HS_5560.B	236	15	83.3	765	5	BM407691
164	15.4	83.3	488	2	BP657167	OY2_21_F0	237	15	83.3	772	9	CC679148
165	15.4	83.3	489	4	BU004187	BU004187	238	15	83.3	778	8	CC093829
166	15.4	83.3	492	6	CB260357	S1-E9624-	239	15	83.3	784	9	CC450880
167	15.4	83.3	494	1	AA854676	AJ63106.s	240	15	83.3	790	5	BM113557
168	15.4	83.3	501	7	CN617766	TGBSTYm3	241	15	83.3	793	8	BZ485455
169	15.4	83.3	504	2	AW744896	1G1_384.H	242	15	83.3	794	8	BZ406661
170	15.4	83.3	506	9	CR400606	ArabiDops	243	15	83.3	803	9	CG230159

244	15	83.3	810	8	BH426750	BH426750	BOGHJ20TR	C 317	14.8	82.2	311	6	CF050330	CF050330	OCM10h03.
245	15	83.3	816	9	CG265163	CG265163	OG0FAV1TH	C 318	14.8	82.2	311	6	CF055390	CF055390	OCN5a03.Y
246	15	83.3	822	9	CC117600	CC117600	NDL.3A7.T	C 319	14.8	82.2	312	6	CF085819	CF085819	QAN15a07.
247	15	83.3	834	8	BZ462557	BZ462557	BONEA51TF	C 320	14.8	82.2	312	6	CD986478	CD986478	QAN22c11.
248	15	83.3	836	8	BZ746910	BZ746910	PUCET35TD	C 321	14.8	82.2	312	6	CD986562	CD986562	QAN23d09.
249	15	83.3	838	8	CC335362	CC335362	OCUAF74TV	C 322	14.8	82.2	312	6	CD986589	CD986589	QAN23g04.
250	15	83.3	839	8	BZ466556	BZ466556	BONNV86TR	C 323	14.8	82.2	312	6	CD987138	CD987138	CAN7e11.Y
251	15	83.3	851	8	BH706797	BH706797	BOMV28TF	C 324	14.8	82.2	312	6	CD989529	CD989529	CAT3a12.Y
252	15	83.3	872	3	CNSADG2	CNSADG2	AX814564	C 325	14.8	82.2	313	6	CF033172	CF033172	QCF13d02.
253	15	83.3	877	9	CC658870	CC658870	AXAriadiops	C 326	14.8	82.2	313	6	CA299790	CA299790	SCMCLV103
254	15	83.3	893	9	CC682520	CC682520	OGXHO03TH	C 327	14.8	82.2	314	6	CD165028	CD165028	ML1-.0093T
255	15	83.3	897	9	CC313186	CC313186	OGXRI83TV	C 328	14.8	82.2	315	6	CD997040	CD997040	CB2d2h02.x
256	15	83.3	902	9	CNS0470R	CNS0470R	AL278316	C 329	14.8	82.2	321	5	CD724442	CD724442	MK.5.13.P
257	15	83.3	902	9	CG265171	CG265171	OG0FAV1TV	C 330	14.8	82.2	323	5	BM552601	BM552601	952076C05
258	15	83.3	906	8	BZ786487	BZ786487	PUGDL02TD	C 331	14.8	82.2	325	8	BZ761225	BZ761225	624.4L4.B
259	15	83.3	907	9	CC658861	CC658861	OGCON90TH	C 332	14.8	82.2	330	6	CD987818	CD987818	QAOSd08.Y
260	15	83.3	916	8	CC343433	CC343433	OGPBR37TV	C 333	14.8	82.2	330	6	CD987839	CD987839	QAOSf08.Y
261	15	83.3	928	9	CC6844197	CC6844197	NDL.132K8	C 334	14.8	82.2	330	6	CF046616	CF046616	QCK27h12.
262	15	83.3	938	9	CNS01TVG	CNS01TVG	AL167029	C 335	14.8	82.2	330	6	CF049903	CF049903	QCL4h12.Y
263	15	83.3	999	6	CH213451	CH213451	SCGGSB114	C 336	14.8	82.2	330	6	CF051213	CF051213	OCM22f10.
264	15	83.3	1020	5	CNS072YK	CNS072YK	clone.BA0	C 337	14.8	82.2	330	6	CF061267	CF061267	OCM19f06.
265	15	83.3	1022	5	BU398891	BU398891	603534208	C 338	14.8	82.2	330	6	CF061497	CF061497	OCT21e02.
266	15	83.3	1101	9	CNS060U2	CNS060U2	AL408861	C 339	14.8	82.2	330	6	CF061902	CF061902	QCT2f06.Y
267	15	83.3	1113	9	CP695295	CP695295	T3 end of	C 340	14.8	82.2	330	6	CF062920	CF062920	OCU15c08.
268	15	83.3	1515	3	CP756660	CP756660	DSAF1.7_A	C 341	14.8	82.2	331	6	CD431949	CD431949	ETH1.12.B
269	14.8	82.2	141	7	BX290873	BX290873	ArabiIdops	C 342	14.8	82.2	331	6	CD987954	CD987954	OA07a10.Y
270	14.8	82.2	147	7	AI901401	AI901401	618003D04	C 343	14.8	82.2	332	6	CF039400	CF039400	QCH36a12.
271	14.8	82.2	155	7	CO296379	CO296379	EX213520.	C 344	14.8	82.2	332	6	CF057339	CF057339	OC031d05.
272	14.8	82.2	156	7	AM017653	AM017653	614061D08	C 345	14.8	82.2	332	6	CF062957	CF062957	QCT3c11.Y
273	14.8	82.2	173	2	CG319258	CG319258	OG1CT74TH	C 346	14.8	82.2	332	6	CF062967	CF062967	OCU15c09.
274	14.8	82.2	192	9	AV931909	AV931909	AV931909	C 347	14.8	82.2	332	6	CF064138	CF064138	OCU7f07.Y
275	14.8	82.2	212	1	AE088235	AE088235	Drosophi1	C 348	14.8	82.2	332	6	CF064176	CF064176	OCUB06.Y
276	14.8	82.2	216	6	CB489366	CB489366	omyKrcDh0	C 349	14.8	82.2	332	7	CO096789	CO096789	GR_Ba20E
277	14.8	82.2	245	6	AM055390	AM055390	614080C11	C 350	14.8	82.2	333	9	TA89B06P	TA89B06P	T_Bruce1
278	14.8	82.2	249	9	CC796146	CC796146	SALX.0964	C 351	14.8	82.2	333	6	CA004073	CA004073	HS16H09T
279	14.8	82.2	252	6	CF063520	CF063520	QCU22e05.	C 352	14.8	82.2	336	6	CF050679	CF050679	OCM15h08.
280	14.8	82.2	257	6	CD989883	CD989883	QAU4d01.Y	C 353	14.8	82.2	340	1	AI193189	AI193189	qe41e07.s
281	14.8	82.2	254	5	BM065667	BM065667	Fgr.8.C24	C 354	14.8	82.2	343	1	CB251982	CB251982	3529.1.19
282	14.8	82.2	264	5	CP115617	CP115617	SSES030N2	C 355	14.8	82.2	346	5	BO163294	BO163294	95208bD10
283	14.8	82.2	267	6	BF488015	BF488015	AT23259.5	C 356	14.8	82.2	349	5	CA925169	CA925169	MTU7TL.P1
284	14.8	82.2	276	2	BI431254	BI431254	949066B02	C 357	14.8	82.2	350	6	CF061192	CF061192	614001P06
285	14.8	82.2	279	4	CB038571	CB038571	BM038571	C 358	14.8	82.2	350	6	AI1795703	AI1795703	OC11f501.
286	14.8	82.2	286	6	CB081591	CB081591	hk53h06.G	C 359	14.8	82.2	350	6	CF028682	CF028682	OCU28g02.
287	14.8	82.2	286	6	BI059597	BI059597	MR4-UT005	C 360	14.8	82.2	352	6	BU062203	BU062203	Fgr.1.E04
288	14.8	82.2	290	6	CB488899	CB488899	omyKrcDh0	C 361	14.8	82.2	354	5	BO068118	BO068118	QA09a10.Y
289	14.8	82.2	292	6	CD436372	CD436372	EL01N0352	C 362	14.8	82.2	354	5	AM126723	AM126723	SA14f01.Y
290	14.8	82.2	293	9	CG685418	CG685418	ZMMBBO15	C 363	14.8	82.2	354	6	CD988118	CD988118	OA09a10.Y
291	14.8	82.2	296	9	CG097332	CG097332	PUIFN61TB	C 364	14.8	82.2	355	2	CF086517	CF086517	RA17z1.7
292	14.8	82.2	297	6	CG962772	CG962772	SDR.132.G	C 365	14.8	82.2	356	4	BI724172	BI724172	RA17z1.7
293	14.8	82.2	298	5	BU571931	BU571931	946166D01	C 366	14.8	82.2	359	1	BG517362	BG517362	947061B05
294	14.8	82.2	299	5	CB831663	CB831663	3529.1.10	C 367	14.8	82.2	360	7	D69723	D69723	CEK073C1F
295	14.8	82.2	300	6	CB488899	CB488899	omyKrcDh0	C 368	14.8	82.2	363	7	CO300168	CO300168	EX1177523.
296	14.8	82.2	300	9	CG097336	CG097336	PUIFN61TD	C 369	14.8	82.2	364	5	BP068550	BP068550	1142804.G
297	14.8	82.2	301	1	AV546446	AV546446	AV546446	C 370	14.8	82.2	368	8	BZ777709	BZ777709	1142804.G
298	14.8	82.2	302	5	BM95630	BM95630	952076C05	C 371	14.8	82.2	369	6	CF057498	CF057498	OC032g06.
299	14.8	82.2	302	6	BM987391	BM987391	QA010902.	C 372	14.8	82.2	370	6	CF056439	CF056439	OC015f08.
300	14.8	82.2	307	4	BM985630	BM985630	952076C05	C 373	14.8	82.2	370	6	CF040058	CF040058	RCW1W16_C
301	14.8	82.2	310	8	CC422313	CC422313	PUIHBB22TD	C 374	14.8	82.2	375	6	CD954718	CD954718	SBQ.217.G
302	14.8	82.2	311	6	CD970550	CD970550	QA019C06.	C 375	14.8	82.2	375	6	CD964221	CD964221	SDN.141.G
303	14.8	82.2	311	6	CD985763	CD985763	QAN14d01.	C 376	14.8	82.2	375	6	CD964653	CD964653	SED.6.Gen
304	14.8	82.2	311	6	CD986140	CD986140	QAN19D12.	C 377	14.8	82.2	376	6	CD967026	CD967026	SER.37.Gen
305	14.8	82.2	311	6	CD986276	CD986276	QAN19G1.Y	C 378	14.8	82.2	376	4	BG241792	BG241792	RA17z1.7
306	14.8	82.2	311	6	CD986416	CD986416	QAN21e09.	C 379	14.8	82.2	376	6	CD945524	CD945524	RA17z1.7
307	14.8	82.2	311	6	CD987053	CD987053	QAN6d12.Y	C 380	14.8	82.2	379	9	AG261131	AG261131	Lotus_cor
308	14.8	82.2	311	6	CD987060	CD987060	QAN6e09.Y	C 381	14.8	82.2	378	5	BE554034	BE554034	BP654034
309	14.8	82.2	311	6	CD987465	CD987465	QAN6i005.Y	C 382	14.8	82.2	379	2	BE552696	BE552696	946083H11
310	14.8	82.2	311	6	CD987747	CD987747	QAN6a01.Y	C 383	14.8	82.2	380	6	CF117128	CF117128	FP264.z1
311	14.8	82.2	311	6	CD988423	CD988423	QAP4a03.Y	C 384	14.8	82.2	380	5	BU051404	BU051404	11104330
312	14.8	82.2	311	6	CD988556	CD988556	QAP4i06.Y	C 385	14.8	82.2	383	5	BO133666	BO133666	952080F11
313	14.8	82.2	311	6	CD989599	CD989599	QAT4a11.Y	C 386	14.8	82.2	385	5	BO133666	BO133666	952080F11
314	14.8	82.2	311	6	CD989622	CD989622	QAT4d10.Y	C 387	14.8	82.2	387	6	CF040826	CF040826	OC115a03.
315	14.8	82.2	311	6	CD989965	CD989965	QAV1d12.Y	C 388	14.8	82.2	389	6	CF036817	CF036817	OCG37c07.
316	14.8	82.2	311	6	CF035517	CF035517	QCG36h01.	C 389	14.8	82.2	389	6	CF057774	CF057774	QCO36g09.

C 390	14.8	82.2	390	6	CF057476	QCO32603.	463	14.8	82.2	491	5	CA895507	B0192E06-
C 391	14.8	82.2	394	8	B2420459	1E6804.b	464	14.8	82.2	492	4	BG103996	BG103996
C 392	14.8	82.2	397	7	CF624454	zmrw505.0	465	14.8	82.2	495	4	BG509546	BG509546
C 393	14.8	82.2	399	1	A1944057	614034D09	466	14.8	82.2	496	6	CF023482	CF023482
C 394	14.8	82.2	403	6	CB082250	hK65909.9	467	14.8	82.2	497	7	CK907743	CK907743
C 395	14.8	82.2	408	7	CF757962	DSAP1.25	468	14.8	82.2	498	7	AC253113	AC253113
C 396	14.8	82.2	411	5	B0410784	GA_EB003	469	14.8	82.2	499	4	CM131785	CM131785
C 397	14.8	82.2	412	5	B6692966	BP692966	470	14.8	82.2	499	4	B1510381	B1510381
C 398	14.8	82.2	414	1	AV803597	AV803597	471	14.8	82.2	499	6	CF043032	CF043032
C 399	14.8	82.2	415	1	A1967286	614051G11	472	14.8	82.2	501	1	A1548084	A1548084
C 400	14.8	82.2	415	1	AV807784	AV807784	473	14.8	82.2	501	6	CA898057	CA898057
C 401	14.8	82.2	415	5	B0410783	GA_EB003	474	14.8	82.2	501	6	CF020279	CF020279
C 402	14.8	82.2	417	6	CA225572	SCRLFL06	475	14.8	82.2	502	1	AV780692	AV780692
C 403	14.8	82.2	418	8	B2696538	SP_BA007	476	14.8	82.2	505	4	BM325647	BM325647
C 404	14.8	82.2	420	4	B4595981	BT595981	477	14.8	82.2	505	4	BM338464	BM338464
C 405	14.8	82.2	425	2	B1251189	EST418449	478	14.8	82.2	505	5	B0663743	B0663743
C 406	14.8	82.2	425	6	CB490711	omykrlb10	479	14.8	82.2	505	6	CB334351	CB334351
C 407	14.8	82.2	430	7	CA487063	EST1709.P	480	14.8	82.2	505	6	CB380651	CB380651
C 408	14.8	82.2	431	1	AV540666	AV540666	481	14.8	82.2	505	7	CO325585	CO325585
C 409	14.8	82.2	431	6	AV927254	AV927254	482	14.8	82.2	506	1	A1901384	A1901384
C 410	14.8	82.2	431	6	CD996528	ORCAh02.x	483	14.8	82.2	506	4	BM551605	BM551605
C 411	14.8	82.2	432	2	B4707040	UT-M-BH3-	484	14.8	82.2	506	8	A2769746	A2769746
C 412	14.8	82.2	434	2	AM017490	614059F06	485	14.8	82.2	506	9	CL719019	CL719019
C 413	14.8	82.2	435	2	AM065436	AM065436	486	14.8	82.2	509	2	BF502367	BF502367
C 414	14.8	82.2	435	2	AM179467	618046B03	487	14.8	82.2	509	6	CD725068	CD725068
C 415	14.8	82.2	436	5	BP049179	BP049179	488	14.8	82.2	510	5	BU694536	BU694536
C 416	14.8	82.2	440	5	BE598324	PI1_81.H0	489	14.8	82.2	510	7	CO139242	CO139242
C 417	14.8	82.2	440	5	BE598324	PI1_81.H0	490	14.8	82.2	511	7	CF796943	CF796943
C 418	14.8	82.2	442	7	CO300457	EX778231.	491	14.8	82.2	512	5	BM059375	BM059375
C 419	14.8	82.2	444	4	BM061072	KS01026B0	492	14.8	82.2	512	6	CF019949	CF019949
C 420	14.8	82.2	444	6	CA011441	HT051070	493	14.8	82.2	513	4	BG509559	BG509559
C 421	14.8	82.2	445	4	BJ484585	BJ484585	494	14.8	82.2	513	4	B1192013	B1192013
C 422	14.8	82.2	447	2	BF500733	AT15745.5	495	14.8	82.2	513	5	BP040284	BP040284
C 423	14.8	82.2	449	2	BE325135	946023G07	496	14.8	82.2	513	7	CO192461	CO192461
C 424	14.8	82.2	449	4	BM073165	MEST61-G0	497	14.8	82.2	514	4	BM340599	BM340599
C 425	14.8	82.2	450	5	B0658867	HD03C020	498	14.8	82.2	516	1	AJ344965	AJ344965
C 426	14.8	82.2	450	5	B0658867	F97_2.O16	499	14.8	82.2	516	2	BP277788	BP277788
C 427	14.8	82.2	455	4	BJ451878	BU451878	500	14.8	82.2	517	2	AM331782	AM331782
C 428	14.8	82.2	455	8	CC085614	CSU-K33r.	501	14.8	82.2	517	7	CE634551	CE634551
C 429	14.8	82.2	456	2	AM416853	52586.MAR	502	14.8	82.2	518	7	CO144784	CO144784
C 430	14.8	82.2	456	7	CM107405	EC2CA30A	503	14.8	82.2	519	1	A1629677	A1629677
C 431	14.8	82.2	457	4	BI096490	949017H03	504	14.8	82.2	520	4	BM072942	BM072942
C 432	14.8	82.2	457	7	CO183533	EC25920.5	505	14.8	82.2	521	4	BM072939	BM072939
C 433	14.8	82.2	460	6	BF007488	QB12908.P	506	14.8	82.2	521	8	BZ484234	BZ484234
C 434	14.8	82.2	460	6	CF007488	QB12908.P	507	14.8	82.2	521	9	CG893876	CG893876
C 435	14.8	82.2	461	7	CO284869	EX167134.	508	14.8	82.2	525	5	BO534739	BO534739
C 436	14.8	82.2	461	7	AM282936	LG1_304.D	509	14.8	82.2	525	9	CG211198	CG211198
C 437	14.8	82.2	466	6	CA891815	B0168B07-	510	14.8	82.2	526	5	B0655331	B0655331
C 438	14.8	82.2	467	4	BI174529	OSFR047H7	511	14.8	82.2	526	6	CD724989	CD724989
C 439	14.8	82.2	467	4	BI174529	OSFR047H7	512	14.8	82.2	527	4	BI097741	BI097741
C 440	14.8	82.2	468	5	BQ294342	1091028D0	513	14.8	82.2	528	4	BQ842533	BQ842533
C 441	14.8	82.2	469	6	CD724930	MX_16_24	514	14.8	82.2	528	6	CA005388	CA005388
C 442	14.8	82.2	472	8	BH854138	SALX_0787	515	14.8	82.2	529	1	A1833441	A1833441
C 443	14.8	82.2	473	6	CD997375	QBD6105.X	516	14.8	82.2	529	6	CA890751	CA890751
C 444	14.8	82.2	475	1	AV923725	AV923725	517	14.8	82.2	529	6	CB179706	CB179706
C 445	14.8	82.2	476	6	CB815392	3529.1.73	518	14.8	82.2	530	5	BO143389	BO143389
C 446	14.8	82.2	479	1	A1947693	603026C08	519	14.8	82.2	530	5	BU695934	BU695934
C 447	14.8	82.2	480	5	BU811359	UL837C10	520	14.8	82.2	530	6	CA144010	CA144010
C 448	14.8	82.2	481	5	BE633388	946024E09	521	14.8	82.2	530	6	CB395975	CB395975
C 449	14.8	82.2	481	7	CO140775	EST833446	522	14.8	82.2	531	2	BF487001	BF487001
C 450	14.8	82.2	484	1	A1455485	LD24478.3	523	14.8	82.2	531	6	CF042088	CF042088
C 451	14.8	82.2	484	5	BO539325	MEST614-D	524	14.8	82.2	532	8	AO666342	AO666342
C 452	14.8	82.2	484	7	CE636693	zmrw0.0	525	14.8	82.2	533	4	BG360645	BG360645
C 453	14.8	82.2	484	7	CO135957	EST833628	526	14.8	82.2	533	6	CF005217	CF005217
C 454	14.8	82.2	485	4	AO602487	HS_218.A	527	14.8	82.2	533	6	CB833657	CB833657
C 455	14.8	82.2	485	6	CA006213	HU05D02u	528	14.8	82.2	535	2	AM090922	AM090922
C 456	14.8	82.2	485	7	CO137678	EST833349	529	14.8	82.2	535	2	AM267411	AM267411
C 457	14.8	82.2	487	2	AM061663	660012G11	530	14.8	82.2	537	5	BO103409	BO103409
C 458	14.8	82.2	488	2	BE510791	946055E05	531	14.8	82.2	537	6	CA485001	CA485001
C 459	14.8	82.2	489	2	BE513397	946070E05	532	14.8	82.2	537	6	CF004663	CF004663
C 460	14.8	82.2	489	7	CF796142	893370.MA	533	14.8	82.2	537	7	CK441039	CK441039
C 461	14.8	82.2	490	7	BU063191	F97_2.Y16	534	14.8	82.2	538	6	CB605394	CB605394
C 462	14.8	82.2	491	5	BQ618922	RKOSQ1D0	535	14.8	82.2	539	1	A0243024	A0243024
C 390	14.8	82.2	390	6	CF057476	QCO32603.	463	14.8	82.2	491	5	CA895507	B0192E06-
C 391	14.8	82.2	394	8	B2420459	1E6804.b	464	14.8	82.2	492	4	BG103996	BG103996
C 392	14.8	82.2	397	7	CF624454	zmrw505.0	465	14.8	82.2	495	4	BG509546	BG509546
C 393	14.8	82.2	399	1	A1944057	614034D09	466	14.8	82.2	496	6	CF023482	CF023482
C 394	14.8	82.2	403	6	CB082250	hK65909.9	467	14.8	82.2	497	7	CK907743	CK907743
C 395	14.8	82.2	408	7	CF757962	DSAP1.25	468	14.8	82.2	498	7	AC253113	AC253113
C 396	14.8	82.2	411	5	B0410784	GA_EB003	469	14.8	82.2	499	4	CM131785	CM131785
C 397	14.8	82.2	412	5	B6692966	BP692966	470	14.8	82.2	499	4	B1510381	B1510381
C 398	14.8	82.2	414	1	AV803597	AV803597	471	14.8	82.2	499	6	CF043032	CF043032
C 399	14.8	82.2	415	1	A1967286	614051G11	472	14.8	82.2	501	1	A1548084	A1548084
C 400	14.8	82.2	415	5	AV807784	AV807784	473	14.8	82.2	501	6	CA898057	CA898057
C 401	14.8	82.2	415	5	B0410783	GA_EB003	474	14.8	82.2	501	6	CF020279	CF020279
C 402	14.8	82.2	417	6	CA225572	SCRLFL06	475	14.8	82.2	502	1	AV780692	AV780692
C 403	14.8	82.2	418	8	B2696538	SP_BA007	476	14.8	82.2	505	4	BM325647	BM325647
C 404	14.8	82.2	420	4	B4595981	BT595981	477	14.8	82.2	505	4	BM338464	BM338464
C 405	14.8	82.2	425	2	B1251189	EST418449	478	14.8	82.2	505	5	B0663743	B0663743
C 406	14.8	82.2	425	6	CB490711	omykrlb10	479	14.8	82.2	505	6	CB334351	CB334351
C 407	14.8	82.2	430	7	CA487063	EST1709.P	480	14.8	82.2	505	6	CB380651	CB380651
C 408	14.8	82.2	431	1	AV540666	AV540666	481	14.8	82.2	505	7</		

682	14.8	82.2	622	6	CF004596	QBH5A07.X	755	14.8	82.2	666	5	BU065758	Fgr_8_G22
683	14.8	82.2	623	6	B5599343	Pt1_88_D0	756	14.8	82.2	667	7	CK755616	CK755616 pat01-2ms
684	14.8	82.2	623	6	CA282084	SCAGSD204	757	14.8	82.2	667	7	CF624183	CF624183 zmrw05.0
685	14.8	82.2	623	6	CA989931	3529.1_11	758	14.8	82.2	667	7	CF630966	CF630966 zmrw04.0
686	14.8	82.2	623	6	CO049056	db05f05.	759	14.8	82.2	668	4	BI487191	BI487191 RE70710-5
687	14.8	82.2	623	8	BI669551	BOMLE94TF	760	14.8	82.2	668	4	BU065958	BU065958 Fgr_8_P13
688	14.8	82.2	623	8	CL700581	SP_BA006	761	14.8	82.2	668	7	CK134651	CK134651 RE70710.3
689	14.8	82.2	625	4	BM332708	MEST176-E	762	14.8	82.2	669	9	BI959787	BI959787 HVSMT002
690	14.8	82.2	625	4	BU079705	946146G08	763	14.8	82.2	669	9	CG377824	CG377824 CG377824
691	14.8	82.2	625	6	CF007122	QBH25G05.	764	14.8	82.2	670	2	BF251484	BF251484 EST148745
692	14.8	82.2	626	6	CF013395	QBH4C04.X	765	14.8	82.2	670	4	BC842553	BC842553 MEST33-C0
693	14.8	82.2	626	7	CK133793	RE38475.3	766	14.8	82.2	670	9	CR487848	CR487848 Medicago
694	14.8	82.2	627	1	AL731320	BNLGH1921	767	14.8	82.2	671	5	BU060768	BU060768 Fgr-C_1_J
695	14.8	82.2	627	6	CF004942	QBH7C11.X	768	14.8	82.2	671	9	CG118904	CG118904 PUKSG6TB
696	14.8	82.2	628	6	CF003636	CF003636 QBH21F06.	769	14.8	82.2	672	4	BC842564	BC842564 MEST33-D0
697	14.8	82.2	629	1	AI388100	GHI1883.5	770	14.8	82.2	672	7	CF632846	CF632846 zmrw04.0
698	14.8	82.2	629	6	CA112029	SCCCLB102	771	14.8	82.2	673	5	BU062831	BU062831 Fgr_2_E24
699	14.8	82.2	629	6	CA486743	WHE4335.E	772	14.8	82.2	674	2	BF251401	BF251401 EST148662
700	14.8	82.2	629	6	CG079821	hp77C10.D	773	14.8	82.2	674	7	CM135442	CM135442 OX1_32_G0
701	14.8	82.2	632	6	CG688195	CHEST-10-F	774	14.8	82.2	676	4	BI577718	BI577718 RE7144T.5
702	14.8	82.2	632	7	CK661212	LP20456.5	775	14.8	82.2	676	4	BM334216	BM334216 MEST134-B
703	14.8	82.2	633	5	BO291309	NKR058.E	776	14.8	82.2	677	1	BU699234	BU699234 LL21n1269
704	14.8	82.2	634	6	CD460615	F909_01d1	777	14.8	82.2	678	1	AI650105	AI650105 AEMTAM85
705	14.8	82.2	635	2	BB625822	BB625822	778	14.8	82.2	678	6	CD462423	CD462423 ETH1_37.F
706	14.8	82.2	635	5	BU042766	PR_LB001	779	14.8	82.2	678	7	CK978064	CK978064 4109386.B
707	14.8	82.2	635	6	CB171494	PK0602601	780	14.8	82.2	679	5	CA806516	CA806516 ESG013d.E
708	14.8	82.2	636	2	AM288963	660019G02	781	14.8	82.2	679	6	CE489928	CE489928 tigr-gss-
709	14.8	82.2	637	1	AI657528	AEMTBD32	782	14.8	82.2	681	7	CO521424	CO521424 3530_1_14
710	14.8	82.2	637	4	BM329776	PIC1_43.F	783	14.8	82.2	681	7	CO530705	CO530705 3530_1_20
711	14.8	82.2	638	9	CB630673	tigr-gss-	784	14.8	82.2	683	5	BM683932	BM683932
712	14.8	82.2	639	7	CB394545	601310555	785	14.8	82.2	683	6	CA236267	CA236267 SCWCFLE501
713	14.8	82.2	639	7	CK369883	zmrw0485	786	14.8	82.2	683	7	CK973943	CK973943 4104695.B
714	14.8	82.2	640	6	CA229997	SCUPFL3C0	787	14.8	82.2	683	9	AG182221	AG182221 Pan. tncg1
715	14.8	82.2	641	1	AI728488	BNLGH1108	788	14.8	82.2	685	5	BM704685	BM704685 BX704685
716	14.8	82.2	641	2	BF488088	AT23172.5	789	14.8	82.2	685	6	CA261088	CA261088
717	14.8	82.2	641	5	CD526392	EST1821.Z	790	14.8	82.2	686	1	AIU68613	AIU68613
718	14.8	82.2	642	5	BO410976	GA_ED003	791	14.8	82.2	686	4	BU581635	BU581635
719	14.8	82.2	642	9	CL703061	SP_BA009	792	14.8	82.2	687	4	BM338300	BM338300 MEST224-SQ
720	14.8	82.2	643	8	AG953905	2M0219B21	793	14.8	82.2	687	5	BO765149	BO765149 EBX008.SQ
721	14.8	82.2	643	8	BB642544	MEST33-B0	794	14.8	82.2	688	4	BM668080	BM668080 MEST375-D
722	14.8	82.2	644	4	BI214977	RE21150.5	795	14.8	82.2	688	2	BF276172	BF276172 GA_ED002
723	14.8	82.2	645	4	BM337075	MEST202-H	796	14.8	82.2	689	2	BM338765	BM338765 MEST231-H
724	14.8	82.2	645	6	CA298342	SCRFPL803	797	14.8	82.2	690	4	BM351687	BM351687 MEST344-G
725	14.8	82.2	645	7	CN148301	WOUNDI.55	798	14.8	82.2	690	4	BC683748	BC683748 Zml0_10e0
726	14.8	82.2	646	5	BX878932	BX878932	799	14.8	82.2	691	4	CA282830	CA282830 SCUPFL301
727	14.8	82.2	647	4	BI579546	RE73592.5	800	14.8	82.2	691	6	BI618277	BI618277
728	14.8	82.2	647	5	BU060077	Fgr-C_0.K	801	14.8	82.2	692	4	BM065774	BM065774
729	14.8	82.2	649	9	BU065702	Fgr_8_E13	802	14.8	82.2	692	5	CA290134	CA290134
730	14.8	82.2	649	9	CB613747	tigr-gss-	803	14.8	82.2	692	6	BU711889	BU711889
731	14.8	82.2	650	2	AM120344	614089C12	804	14.8	82.2	693	4	CA289543	CA289543 SCAGFL800
732	14.8	82.2	650	9	CL785314	OR_BBA010	805	14.8	82.2	694	6	CA202228	CA202228 SCBFL103
733	14.8	82.2	651	4	BM266381	MEST395-F	806	14.8	82.2	694	6	CF669689	CF669689 RTGNT1_45
734	14.8	82.2	651	5	BU064549	Fgr_5_C14	807	14.8	82.2	694	7	BC936617	BC936617 Zm08_04b0
735	14.8	82.2	652	7	CN135371	OX1_32_G0	808	14.8	82.2	695	4	CA094351	CA094351 SCCCLB100
736	14.8	82.2	654	8	BB677854	PUBBGS7TD	809	14.8	82.2	695	6	AG182317	AG182317 Pan. tncg1
737	14.8	82.2	655	7	CO530706	3530_1_20	810	14.8	82.2	695	9	BM446346	BM446346 WOUNDI.56
738	14.8	82.2	656	5	BU062210	Fgr_1_H21	811	14.8	82.2	696	6	CN148382	CN148382
739	14.8	82.2	656	7	CN881248	010425AAS	812	14.8	82.2	696	7	BM274387	BM274387 CH230-35D
740	14.8	82.2	657	6	CD208455	H51_37_C1	813	14.8	82.2	697	8	BM430764	BM430764 MEST325-F
741	14.8	82.2	658	4	BB842131	MEST33-C0	814	14.8	82.2	697	4	BM269384	BM269384 ESTB14334
742	14.8	82.2	658	4	BB842131	MEST33-E1	815	14.8	82.2	697	7	BU065866	BU065866 Fgr_8_L12
743	14.8	82.2	659	4	BM074759	MEST297-C	816	14.8	82.2	698	5	CK761162	CK761162 act102-5ms
744	14.8	82.2	660	4	BB874190	MEST147-D0	817	14.8	82.2	698	7	CF273034	CF273034 EST19596.Z
745	14.8	82.2	660	6	CB486596	omykrcbho	818	14.8	82.2	699	6	CF006279	CF006279
746	14.8	82.2	660	6	CL827350	OR_CBA004	819	14.8	82.2	700	6	BM269384	BM269384 MEST325-B
747	14.8	82.2	661	5	BU064367	Fgr_4_K22	820	14.8	82.2	701	4	BM430764	BM430764 MEST325-B
748	14.8	82.2	661	5	CA243864	SCQFL407	821	14.8	82.2	701	4	BM430764	BM430764 MEST325-B
749	14.8	82.2	663	4	BM348162	MEST287-C	822	14.8	82.2	702	6	CA111363	CA111363 SCAGFL806
750	14.8	82.2	663	5	BB623048	USDA-FP-0	823	14.8	82.2	704	7	CO521425	CO521425 3530_1_14
751	14.8	82.2	663	5	BO064289	Fgr_4_H15	824	14.8	82.2	704	9	CL828295	CL828295 OR_CBA004
752	14.8	82.2	663	6	CA385174	666023 NC	825	14.8	82.2	705	5	BU066073	BU066073 Fgr_9_E12
753	14.8	82.2	663	6	CD208359	H51_37_C1	826	14.8	82.2	705	6	CF006594	CF006594
754	14.8	82.2	664	8	BB67693	CH240_224	827	14.8	82.2	706	5	BM312416	BM312416

974 14.8 82.2 859 7 CF824787 EST702169
C 975 14.8 82.2 863 7 C0030855 EST809239
C 976 14.8 82.2 864 7 C0014376 EST784758
C 977 14.8 82.2 872 2 CG263606 GCMUK377H
C 978 14.8 82.2 873 2 BR303626
C 979 14.8 82.2 873 7 CF819514 EST696896
C 980 14.8 82.2 873 9 CL687900 PRI0148a
C 981 14.8 82.2 875 9 CG302499 CG3CB53TV
C 982 14.8 82.2 876 7 C0008648 EST796983
C 983 14.8 82.2 876 9 C0005527 EST793862
C 984 14.8 82.2 877 7 C0005527 EST793862
C 985 14.8 82.2 878 3 AY110987 Zee may
C 986 14.8 82.2 881 8 A2128036
C 987 14.8 82.2 882 7 C0010892
C 988 14.8 82.2 887 2 BE565682 60133243
C 989 14.8 82.2 889 9 CG377810 CG3DV77TH
C 990 14.8 82.2 893 6 CD458627 FG08.1212
C 991 14.8 82.2 895 9 CG035407 PUJFAL2TD
C 992 14.8 82.2 895 9 CG383781 CG2C117TH
C 993 14.8 82.2 896 1 A1066632 mgea0003c
C 994 14.8 82.2 896 6 C0025150 UCRHY8.0
C 995 14.8 82.2 896 7 C0025150 EST80534
C 996 14.8 82.2 898 1 A1666471 AL666471
C 997 14.8 82.2 900 5 B0618940 RNOSEQ1F0
C 998 14.8 82.2 900 5 B0618946 RNOSEQ1F0
C 999 14.8 82.2 900 5 B0619165 RNOSEQ4E0
C 1000 14.8 82.2 900 5 B0619196 RNOSEQ4H0

ALIGNMENTS

RESULT 1 164 bp DNA linear GSS 02-OCT-2003
LOCUS CG665075
DEFINITION OST178045 Mus musculus 129Sv/Ev Mus musculus genomic clone
ACCESSION CG665075
VERSION CG665075.1 GI:37488924
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 164)
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J., Piggett,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W., Jr., Kipp,P., Kohhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,C. and Sands,A.T.

TITLE Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES Location/Qualifiers

source 1..164
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST178045"
/cell_type="embryonic stem cell"
/clone_1ib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGACACACAGATTTCG 17
DB 14 GAGACACACAGATTTCG 30

RESULT 2 383 bp DNA linear GSS 01-OCT-2003
LOCUS CG559318
DEFINITION OST178045 Mus musculus 129Sv/Ev Mus musculus genomic clone
ACCESSION CG559318
VERSION CG559318.1 GI:37345905
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 383)
AUTHORS Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J., Piggett,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A., Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C., Key,B.W., Jr., Kipp,P., Kohhauff,B., Ma,Z.-Q., Markesich,D., Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z., Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N., Zhu,Q., Person,C. and Sands,A.T.

TITLE Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP
OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES Location/Qualifiers

source 1..383
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST178045"
/cell_type="embryonic stem cell"
/clone_1ib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 383;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTCG 17
DB 14 GAGACACACAGATTTCG 30

RESULT 3 936 bp DNA linear GSS 07-JUL-2001
LOCUS CNS07414
DEFINITION Clone BA0AB024E03 of library BA0AB from strain CLB 210 of Kluveromyces lactic, genomic survey sequence.
ACCESSION AL428222
VERSION AL428222.1 GI:12211416
KEYWORDS GSS.

SOURCE Kluveromyces lactic
ORGANISM Kluveromyces lactic

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Kluveromyces.


```

REFERENCE
AUTHORS      1 (bases 1 to 936)
              Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
              Bolotin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
              de-Montigny,J., Dujon,B., Durren,P., Lepingle,A., Llorente,B.,
              Malpertuy,A., Neuvéglise,C., Olier-Kalogeropoulos,O., Potier,S.,
              Saurin,M., Tekala,F., Toffano-Nioche,C., Weslowski-Louvel,M.,
              Winkler,P. and Weissenbach,J.
TITLE        Genomic exploration of the hemiascomycetous yeasts : 1. A set of
              yeast species for molecular evolution studies
JOURNAL      FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE      20584711
PUBMED       1152876
AUTHORS      2 (bases 1 to 936)
              Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
              Duchateau-Ngyen,G., Lemaire,M., Marmelisse,R., Montrocher,R.,
              Robert,C., Termier,M., Winkler,P. and Weslowski-Louvel,M.
TITLE        Genomic exploration of the hemiascomycetous yeasts : 11.
JOURNAL      Kluweromyces lactis
MEDLINE      FEBS Lett. 487 (1), 66-70 (2000)
PUBMED       1152886
AUTHORS      3 (bases 1 to 936)
              Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
              2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
              seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      This GSS is part of a random genomic sequencing program of thirteen
              yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
              exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
              Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
              lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
              angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
              Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
              5 kb were prepared and both extremities were sequenced. See
              keywords for description of this sequence and for the sequence of
              the other extremity of this insert.
FEATURES
source       Location/Qualifiers
              1..936
               /organism="Kluyveromyces lactis"
               /mol_type="Genomic DNA"
               /strain="CIB 210"
               /variety="lactis"
               /db_xref="taxon:28985"
               /clone="BA0AB024EB3"
               /clone_11b="BA0AB"
               complement(<463..>.936)
               /note="similar to Saccharomyces cerevisiae ORF YD1091c [
               weak similarity to mouse FAF1 protein ]"
               /evidence=not_experimental
ORIGIN
Query Match          94.4%; Score 17; DB 9; Length 936;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GAGATCAACAAGATTG 17
         |||
Db       875 GAGACACACAAAGATTGC 859
RESULT 4
LOCUS     BG410386/c                179 bp    mRNA    linear    EST 13-WAR-2001
DEFINITION sequence.
ACCESSION BG410386
VERSION   BG410386.1 GI:13315834
KEYWORDS  EST.
SOURCE    Zea mays
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

```

REFERENCE	Clade: panicoidae; Andropogoneae; Zea.
AUTHORS	1 (bases 1 to 179)
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	University
COMMENT	Unpublished (1999) Contact: Walbot, V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 947045 row: H column: 08. Location/Qualifiers
FEATURES	1..179
source	/organism="Zea mays" /mol_type="mRNA" /cultiivar="B73" /db_xref="taxon:4577" /tissue_type="leaf and stem, including leaf base" /dev_stage="2 week old seedling (3 leaves)" /lab_host="XJ1-Blue" /clone_lbp="947 - 2 week shoot from Baykan lab" /note="Organ: shoot; Vector: Lambda ZAP (pbiscript SK-); Site 1: EcorI, Site 2: XhoI; Directionally cloned using Stratagene's UniZap XR cDNA cloning kit with the 5' end at the EcocRI site. The library represents 8 x 10es independent recombinant phage. The plants were greenhouse grown."
ORIGIN	
Query Match	91.1%; Score 16.4; DB 4; Length 179;
Best Local Similarity	94.4%; Pred. No. 1.6e+03;
Matches	17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 GAGAACCAACAAGATTCCG 18
Db	34 GAGAACATCAAGATTCCG 17
RESULT 5	
LOCUS	CD460790 357 bp mRNA linear EST 14-JUN-2004
DEFINITION	Fg09_01n04_A Fg09 AAFc EOCRC Fusarium graminearum simple_substrate
GI	Gibberella zeae cDNA clone Fg09_01n04, mRNA sequence.
ACCESSION	CD460790
VERSION	CD460790.2 GI:48688810
KEYWORDS	EST.
SOURCE	Gibberella zeae
ORGANISM	Gibberella zeae Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocerales; Nectriaceae; Gibberella.
REFERENCE	1 (bases 1 to 357)
AUTHORS	Watson,R.J., Heyes,R., Couroux,P., De Moors,A., Harris,L.J., Hattori,U., Lacroix,C., Masotti,M., Queller,T., Robert,L.S., Singh,J.A., Sprott,D. and Tinker,N.A.
TITLE	A cDNA library prepared from Fusarium graminearum grown on a simple substrate
JOURNAL	Unpublished (2003)
COMMENT	On Jun 3, 2003 this sequence version replaced gi:31375530. Contact: Watson, Robert.J. Eastern Cereal and Oilseed Research Centre Agriculture and Agri-Food Canada Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA Tel: (613) 759-1655 Fax: (613) 759-1701 Email: watsonrj@agr.gc.ca.
FEATURES	1..357
source	/organism="Gibberella zeae" /mol_type="mRNA"

/strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg09_01n04"
 /tissue_type="Mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /clone_lib="Fg09_AAFc_ECORC_Fusarium graminearum simple_substrata"
 /note="Vector: pBluescript II+, site 1: EcorI; site 2: XhoI; Fusarium graminearum grown on a simple substrate--minimal media supplemented with amino acids."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 357;
 Best Local Similarity 94.4%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGAACACACAGATTGCG 18
 |||||
 223 GAGAACACACAGATTGCG 240

RESULT 6
 A0210247 481 bp DNA linear GSS 18-SEP-1998
 LOCUS HS_3115_B2_C12_T7 CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens_genomic_clone Plate=3115 Col=24 Row=F, genomic survey
 sequence.

ACCESSION A0210247
 VERSION A0210247.1 GI:3619216
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 481)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 3115 row: F column: 24
 Class: BAC ends
 High quality sequence stop: 481.

FEATURES

Location/Qualifiers
 1..481
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=3115 Col=24 Row=F"
 /sex="male"
 /note="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBlotBAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 481;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGAACACACAGATTGCG 18
 |||||
 |||||

Db 242 GAGAACACACAGATTGCG 259

RESULT 7
 CD657925/c 504 bp mRNA linear EST 19-JUN-2003
 LOCUS EESTEF5909.y1 Eimeria tenella M5-6 Excised cDNA Eimeria tenella
 DEFINITION CDNA 5', mRNA sequence.

ACCESSION CD657925
 VERSION CD657925.1 GI:3190054
 KEYWORDS EST.
 SOURCE Eimeria tenella
 ORGANISM Eimeria tenella
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.

REFERENCE

1 (bases 1 to 504)
 Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Stepec,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schunk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
 WashU-Merck Eimeria tenella project
 Unpublished (1999)
 CONTACT: David Sibley, Ph.D.
 WashU-Merck Eimeria tenella project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 CONTACT David Sibley (toxest@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco.

FEATURES

Location/Qualifiers
 1..504
 /organism="Eimeria tenella"
 /mol_type="mRNA"
 /db_xref="taxon:5802"
 /dev_stage="IS18"
 /lab_host="E. coli DH10B (GeneHog, Invitrogen, Inc)"
 /clone_lib="Eimeria tenella M5-6 Excised cDNA"
 /note="Vector: pBluescript SK-; Site_1: EcorI; Site_2: XhoI; Sporozoites were obtained from in vitro sporulated and excysted oocysts of E. tenella grown in chickens. cDNA was synthesized from poly RNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcorI adapters were ligated to the cDNA and products were size-selected on sephacryl S500. The cDNA were ligated to EcorI/XhoI prepared lambda ZapII (Stratagene). The primary library was mass excised using Exsist helper phage (Stratagene). The phagemids were precipitated with PBG 8000, extracted with phenol/chloroform and electroporated into DH10B cells. The library may contain a small percentage of host or bacterial contaminants. Library materials provided by: Paul Liberator, Merck Research Labs Library constructed by: Paul Liberator, Merck Research Labs."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 504;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGAACACACAGATTGCG 18
 |||||
 347 GAGAACACACAGATTGCG 330

RESULT 8
 BM076070/c 514 bp mRNA linear EST 13-NOV-2001
 LOCUS BM076070
 DEFINITION MEST364-D04.T3 ISUM5-RN Zea mays CDNA clone MEST364-D04 3', mRNA

sequence.
 BMO76070.1 GI:16923078
 VERSION BMO76070.1
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 514)
 Wen, T.J., Qiu, F., Guo, L., Ashlock, D.A. and Schnable, P.S.
 Expressed Sequence Tags from B73 Maize: various stages and tissues including seedlings treated with a variety of hormones
 Unpublished (2001)
 CONTACT: Patrick S. Schnable
 Schnable Laboratory
 Iowa State University
 G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
 Tel: 515-294-0975
 Fax: 515-294-2239
 Email: schnable@iastate.edu
 Individual basecall and confidence value were assigned using the Phred software,
 (<http://depts.washington.edu/ventures/collabtr/direct/index.htm>#rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (<http://www.tigr.org/softlab/>). Lucy parameters were set to ensure an overall trimmed quality of 97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.
 PCR Primers
 FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA CTA TGG)
 BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC TAA AG)
 Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
 Location/Qualifiers
 1..514
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="MEST364-D04"
 /tissue_type="mixed"
 /lab_host="DH10B"
 /clone_lib="RISUM5-RN"
 /note="Vector: pRTT3PAC; Site 1: EcoRI; Site 2: NotI; Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG), Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels (3, 5, 10, 15, 20, 25, 30, DAG), Adventitious roots (65 DAG), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0 cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated first ear, ear shank, etiolated seedlings, callus, Cycloheximide-treated callus, Anaerobic treated seedlings, NAA (6-Naphthalene acetic acid)-treated seedlings, Kinetin-treated seedlings, ACPG (1-aminocyclopropane-1-carboxylic acid)-treated seedlings, Brassinolide-treated seedlings, ABA (Abscisic acid)-treated seedlings, GA (Gibberellic acid)-treated seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5'-AACTGAGAGATCGCGCGAGAAATTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pRTT3PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome

ORIGIN Research 6: 791-806, 1996)."
 Query Match 91.1%; Score 16.4; DB 4; Length 514;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 281 GAGAACATCAAGATTCGC 264
 1 GAGAACATCAAGATTCGC 18
 |||||
 281 GAGAACATCAAGATTCGC 264
 RESULT 9
 LOCUS CO817404
 DEFINITION 24h after treatment with salicylic acid Fragaria x ananassa cDNA
 ACCESSION CO817404.1 GI:51048151
 VERSION CO817404
 KEYWORDS EST.
 SOURCE Fragaria x ananassa
 ORGANISM Fragaria x ananassa
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosids; eustosids I; Rosales; Rosaceae; Rosoideae; Fragaria.
 1 (bases 1 to 768)
 Folta, K., Stewart, P. and Bies, D.
 Development of an EST Database for Octoploid Strawberry (Fragaria x ananassa, cv 'Strawberry Festival')
 Unpublished (2004)
 CONTACT: Folta, K
 Horticultural Sciences Department, Plant Molecular and Cellular Biology Program
 University of Florida
 1301 Field Hall, PO BOX 110690, Gainesville, FL 32611, USA
 Tel: 352 392 3814
 Fax: 352 392 5653
 Email: kfolta@ufl.edu
 Total High Quality Bases = 637
 Seq primer: CACAGGAACAGCTATGACC
 High quality sequence stop: 768.
 Location/Qualifiers
 1..768
 /organism="Fragaria x ananassa"
 /mol_type="mRNA"
 /cultivar="Strawberry Festival"
 /db_xref="taxon:3747"
 /clone="FA_SBA0011C05r"
 /tissue_type="whole plant treated 24 h with 1mM salicylic acid (spray), 4 mM soak"
 /lab_host="E. coli"
 /clone_lib="Fragaria x ananassa 'Strawberry Festival', cultivar 24h after treatment with salicylic acid"
 /note="Vector: pBluescript; Site 1: EcoRI; Site 2: XhoI; Fragaria x ananassa 'Strawberry Festival' cultivar (octoploid), total RNA derived from whole plant vegetative tissues 24 h after treatment with salicylic acid (4 uM foliar spray and 1 uM root drench). For more details on library preparation and sequence analysis see http://www.hos.ufl.edu/shortcake and http://www.genome.clemson.edu/gtr To order clones from this library see http://www.arabidopsis.thaliana.com/strawberry/clones.htm"
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 7; Length 768;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 326 GAGAAAACCAAGATTCGC 343
 1 GAGAACATCAAGATTCGC 18
 |||||
 326 GAGAAAACCAAGATTCGC 343

RESULT 10
 BI823206/c
 LOCUS
 DEFINITION
 BI823206 1138 bp mRNA linear EST 04-OCT-2001
 6030394691 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180272 5',
 mRNA sequence.
 ACCESSION
 BI823206
 VERSION
 BI823206.1 GI:15934756
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1138)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
 Plate: LIML149 row: j column: 17
 High quality sequence stop: 324.
 Location/Qualifiers
 1..1138
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5180272"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 PCW-SORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 4; Length 1138;
 Best Local Similarity 94.4%; Pred. No. 1.9e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGAACAAAGATTGCG 18
 |||||
 DB 1011 GAGAACAAAGATTGCG 994
 |||||
 RESULT 11
 BE574801
 LOCUS
 DEFINITION
 BE574801 368 bp mRNA linear EST 07-DEC-2001
 H19 Triphysaria versicolor root-tip, early DMBQ-induced transcript
 cDNA library Triphysaria versicolor cDNA, mRNA sequence.
 ACCESSION
 BE574801
 VERSION
 BE574801.1 GI:12001131
 KEYWORDS
 EST.
 SOURCE
 Triphysaria versicolor
 ORGANISM
 Triphysaria versicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Lamiales; Orobanchaceae; Rhinanthaceae;
 Triphysaria.
 REFERENCE
 1 (bases 1 to 368)

AUTHORS
 TITLE
 Matvienko, M., Torres, M.J. and Yoder, J.I.
 Transcriptional responses in the hemiparasitic plant Triphysaria
 versicolor to host plant signals
 JOURNAL
 Plant Physiol. 127 (1), 272-282 (2001)
 MEDLINE
 21437952
 PUBMED
 11533755
 COMMENT
 Contact: Yoder, J.I.
 John I. Yoder Research Lab, Dept. of Vegetable Crops
 University of California at Davis
 137 Asmunden Hall, One Shields Drive, Davis, CA 95616, USA
 Tel: 530 752 1741
 Fax: 530 752 9659
 Email: jyoder@ucdavis.edu
 length = 368 bp.
 Location/Qualifiers
 1..368
 /organism="Triphysaria versicolor"
 /mol_type="mRNA"
 /db_xref="taxon:64093"
 /tissue_type="root-tips ~5mm length"
 /dev_stage="3-4 weeks growth"
 /lab_host="E. coli"
 /clone_id="Triphysaria versicolor root-tip, early
 DMBQ-induced transcript cDNA library"
 /note="Vector: PCR2.1 TA Cloning System, Invitrogen,
 Carlsbad, CA; PCR-based suppression subtractive
 hybridization cDNA library"
 ORIGIN
 Query Match 88.9%; Score 16; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GACACAAAGATTGCG 18
 |||||
 DB 86 GACACAAAGATTGCG 101
 |||||
 RESULT 12
 CE085246
 LOCUS
 DEFINITION
 CE085246 426 bp DNA linear GSS 24-SEP-2003
 tigr-gss-dog-17000359005989 Dog Library Canis familiaris genomic,
 genomic survey sequence.
 ACCESSION
 CE085246
 VERSION
 CE085246.1 GI:35152092
 KEYWORDS
 GSS.
 SOURCE
 Canis familiaris (dog)
 ORGANISM
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 REFERENCE
 1 (bases 1 to 426)
 Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Ruch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
 The dog genome: survey sequencing and comparative analysis
 JOURNAL
 Science 301 (5641), 1898-1903 (2003)
 MEDLINE
 22875432
 PUBMED
 14512627
 COMMENT
 Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.
 Location/Qualifiers
 1..426
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_id="Dog Library"

ORIGIN /note="Site 1: Betxi; Libraries were prepared from peripheral blood"

Query Match 88.9%; Score 16; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||
Db 111 GAGAACACACAGATTC 126

RESULT 13 CK130476 646 bp mRNA linear EST 02-DEC-2003
LOCUS CK130476
DEFINITION AT08181.jprtime AT Drosophila melanogaster adult testes P0TB7
Drosophila melanogaster cDNA clone AT08181.3 similar to CG9905:
F8an0009905 GO: [] located on: 4 102D1-102D1; 08/14/2002, mRNA
sequence.

ACCESSION CK130476
VERSION CK130476
KEYWORDS GI:38622412
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 646)
Stapleton, M., Brokstein, P., Hong, L., Agbayan, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guatin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Mitra, S., Mungall, C. J., Nunco, J.,
Palele, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S. E., Celniker, S. and Rubin, G. M.
BDGP/HMT AT Drosophila EST Project
Unpublished (2000)
Other ESTs: AT08181.5prtime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AF003843: arm:4 [434780..751419] estimated-cyto:102c1-102e1:
02/05/2002

Plate: AT 81 row: G column: 9
High quality sequence stop: 560.
Location/Qualifiers

FEATURES
source 1..646
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT08181"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT-10-AT.120; DHS-alpha. Plates
AT.121-AT.319; DHS-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes P0TB7"
/note="Organ: ADULT testes; Vector: P0TB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into P0TB7. Plasmid cDNA library."

ORIGIN

Query Match 88.9%; Score 16; DB 7; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCG 17
|||||
Db 583 AGAACACACAGATTCG 598

RESULT 14 B2178809 690 bp DNA linear GSS 11-OCT-2002
LOCUS B2178809
DEFINITION CH230-443N20.TVB CHORI-230 Segment 2 Rattus norvegicus genomic
clone CH230-443N20, genomic survey sequence.

ACCESSION B2178809
VERSION B2178809
KEYWORDS GI:23822940
SOURCE GSS.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 690)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, B., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C. M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-443N20.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cbo.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_eirng_information.htm). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 443 row: N column: 20
Seq primer: 17
Class: BAC ends.
Location/Qualifiers

FEATURES
source 1..690
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SnHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-443N20"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 690;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||
Db 565 GAGAACACACAGATTC 580

RESULT 15 CC082521 744 bp DNA linear GSS 16-APR-2003
LOCUS CC082521
DEFINITION CSU-K33r.3868.T7 CSU-K33r Aedes aegypti genomic clone
CSU-K33r.3868, genomic survey sequence.

ACCESSION CC082521
VERSION CC082521.1 GI:29931623

KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes; Stegomyia.
REFERENCE 1 (bases 1 to 744)
AUTHORS Loftus, B., Shetty, J., Severnson, D., Brown, S. and Knudson, D.
TITLE End sequencing of Aedes aegypti BACs
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CSU-K33.38G8.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: ena@tigr.org
Library was provided by Susan Brown and Dennis Knudson at Colorado State University.
Seq primer: T7
Class: BAC ends.

FEATURES
source location/Qualifiers
1..744
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Rexville"
/db_xref="taxon:7159"
/clone="CSU-K33.38G8"
/clone_1db="CSU-K33"
/note="Vector: pBelBAC11; Site_1: HindIII"

ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACAACAGATTTC 17
|||||
544 AGAACAACAGATTTC 559

Db

RESULT 16
CL491245/c 892 bp DNA linear GSS 01-APR-2004
LOCUS SAIL_554_A05.v2 SAIL Collection Arabidopsis thaliana genomic clone
DEFINITION SAIL_554_A05.v2, genomic survey sequence.
ACCESSION CL491245
VERSION CL491245.1 GI:45975701
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 892)
AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchinson, D., Kimmery, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
MEDLINE 22356987
PUBMED 12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
SynGene Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS823429; T-DNA left border flanking sequences of SynGene Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not

single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..892
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_554_A05.v2"
/clone_1db="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN
Query Match 88.9%; Score 16; DB 9; Length 892;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTTC 16
|||||
622 GAGAACACAGATTTC 607

Db

RESULT 17
BU017470 227 bp mRNA linear EST 23-AUG-2002
LOCUS OHE15P17.YG.ab1 OH_EFGH sunflower RHA280 Helianthus annuus cDNA
DEFINITION clone OHE15P17, mRNA sequence.
ACCESSION BU017470
VERSION BU017470.1 GI:22452990
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asterales; Asteroideae; Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 227)
AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., Van Damme, M., Lavelle, D., Chevallier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
TITLE Lettuce and Sunflower ESTs from the Composite Genome Project
JOURNAL http://compgenome.ucdavis.edu/
COMMENT Unpublished (2002)
CONTACT: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig192, see http://cgpdb.ucdavis.edu/ for details.
Plate: OHE15 row: P column: 17.

FEATURES
source location/Qualifiers
1..227
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE15P17"
/lab_host="E.coli"
/clone_1db="OH_EFGH sunflower RHA280"
/note="Vector: pBRCDNA51AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformants made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/

ORIGIN TAG_TISSUE=shoots environmental stress
TAG_LIB=OH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

Query Match 85.6%; Score 15.4; DB 5; Length 227;
Best Local Similarity 94.1%; Pred. No. 5.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAAGATTCCG 98

RESULT 18 BU018252 229 bp mRNA linear EST 23-AUG-2002
LOCUS OHE18B18.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
DEFINITION clone OHE18B18, mRNA sequence.

ACCESSION BU018252 GI:22453772
VERSION BU018252.1
KEYWORDS EST:
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 229)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J.,
Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L., and Bradford,K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://compgenome.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aasmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig192, see http://cgpdb.ucdavis.edu/
for details.
Plate: OHE18 row: B column: 18.

FEATURES
source Location/Qualifiers
1..229
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE18B18"
/lab_host="E.coli"
/clone_lib="OH_EFGHJ sunflower RHA280"
/note="Vector: pRCNDNA5flab; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_SEQ=Not found"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 229;
Best Local Similarity 94.1%; Pred. No. 5.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 AGAACACACAGATTCCG 18

Db 81 AGAAAAACAAGATTCCG 97

RESULT 19 BU016050 269 bp mRNA linear EST 23-AUG-2002
LOCUS OHE11P07.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
DEFINITION clone OHE11P07, mRNA sequence.

ACCESSION BU016050 GI:22451570
VERSION BU016050.1
KEYWORDS EST:
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 269)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J.,
Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L., and Bradford,K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://compgenome.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aasmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig192, see http://cgpdb.ucdavis.edu/
for details.
Plate: OHE11 row: P column: 07.

FEATURES
source Location/Qualifiers
1..269
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE11P07"
/lab_host="E.coli"
/clone_lib="OH_EFGHJ sunflower RHA280"
/note="Vector: pRCNDNA5flab; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=OH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 269;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAAGATTCCG 98

RESULT 20 CD944524 270 bp mRNA linear EST 15-JUL-2003
LOCUS RDU 72 Genetagi Zea mays cDNA, mRNA sequence.
DEFINITION CD944524
ACCESSION CD944524 GI:32792288
VERSION

KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 270)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).
Location/Qualifiers

FEATURES
source 1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="GenetAg1"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTG 17
|||
86 GAGACACACATGATTG 102

Db

RESULT 21
CD946927 270 bp mRNA linear EST 15-JUL-2003
LOCUS REV 27 GenetAg1 Zea mays cDNA, mRNA sequence.
DEFINITION CD946927
ACCESSION CD946927.1 GI:32794691
VERSION
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 270)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).
Location/Qualifiers

FEATURES
source 1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="GenetAg1"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTG 17
|||
86 GAGACACACATGATTG 102

Db

RESULT 22
CD967114 270 bp mRNA linear EST 16-JUL-2003
LOCUS SES 132 GenetAg2 Zea mays cDNA, mRNA sequence.
DEFINITION CD967114
ACCESSION CD967114.1 GI:32827436
VERSION
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 270)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).
Location/Qualifiers

FEATURES
source 1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="GenetAg2"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTG 17
|||
86 GAGACACACATGATTG 102

Db

RESULT 23
CD967248 270 bp mRNA linear EST 16-JUL-2003
LOCUS SES 61 GenetAg2 Zea mays cDNA, mRNA sequence.
DEFINITION CD967248
ACCESSION CD967248.1 GI:32827570
VERSION
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 270)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com> and <http://genoplatte-info.infobiogen.fr>).
Location/Qualifiers

FEATURES
source 1..270

ORIGIN

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="Genetags2"

Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTGC 17
185 GAGAACACACAGATTGC 169

RESULT 24
AW977559

LOCUS AW977559 273 bp mRNA linear EST 02-JUN-2000
DEFINITION EST189668 MAGR resequences, MAGO Homo sapiens cDNA, mRNA sequence.
ACCESSION AW977559
VERSION AW977559.1 GI:8168812
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 273)
Hegde, P., Qi, R., Abernathy, K., Dhar, S., Gaspard, R., Gay, C., Holt, I. E., Speed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
CONTACT: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@igir.org
Plate: 380

TITLE
JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..273
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGO"
/note="Vector: pBluescriptSKm"

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 273;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTGC 18
251 AGAACACACAGATTGC 267

RESULT 25
BQ910769

LOCUS BQ910769 284 bp mRNA linear EST 19-AUG-2002
DEFINITION CHA15C10.yg.ab1 CH_ABCDI sunflower RHA801 Helianthus annuus cDNA
clone CHA15C10, mRNA sequence.
ACCESSION BQ910769.1 GI:22309548
VERSION BQ910769
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus

Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;

REFERENCE
AUTHORS
Heliantheae; Helianthus.
1 (bases 1 to 284)
Kozik, A., Michelmore, R. W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M. S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aemundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmoreveg@mail.ucdavis.edu]
Belongs to config CH_CA.Config192, see http://cgdb.ucdavis.edu/
for details.
Plate: CHA15 row: C column: 10.

JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..284
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="CHA15C10"
/lab_host="E.coli"
/note="Vector: pBRCDNA51AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG_ISSUE=shoots environmental stress
TAG_LIB=CH ABCDI sunflower RHA801
TAG_SEQ=TCGCAACGCG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 284;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTGC 18
82 AGAACACACAGATTGC 98

RESULT 26
BU016309

LOCUS BU016309 285 bp mRNA linear EST 23-AUG-2002
DEFINITION CHB12L17.yg.ab1 CH_ERGHJ sunflower RHA280 Helianthus annuus cDNA
clone CHB12L17, mRNA sequence.
ACCESSION BU016309.1 GI:22451829
VERSION BU016309
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus

Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 285)
Kozik, A., Michelmore, R. W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M. S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)

JOURNAL
TITLE

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Ammundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig192, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHE12 row: L column: 17.

FEATURES

source

Location/Qualifiers

1..285
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QH12L17"
/lab_host="E.coli"
/clone_id="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCNAST1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCACGGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 285;
Best Local Similarity 94.1%; Pred.No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAGATTCCG 98

Db

RESULT 27

LOCUS

BH021781 286 bp mRNA linear EST 23-AUG-2002

DEFINITION

QH3407.Y9.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA

ACCESSION

BH021781 GI:22457301

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Helianthus annuus (common sunflower)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 286)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Liu, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Stabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lecture and Sunflower ESTs from the Compositae Genome Project
http://composgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Ammundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig192, see http://cgpdb.ucdavis.edu/
for details.

Plate: QHE4 row: e column: 07.

FEATURES

source

1..286
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QH4E07"
/lab_host="E.coli"
/clone_id="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCNAST1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCACGGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 286;
Best Local Similarity 94.1%; Pred.No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAGATTCCG 98

Db

RESULT 28

LOCUS

CA485884 300 bp mRNA linear EST 14-NOV-2002

DEFINITION

WHE4324_C12_F2425 wheat meiotic anther cDNA library Triticum

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Triticum aestivum (bread wheat)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 300)
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Izzo, G.R.,
Pham, J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
Unpublished (2002)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 510595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer
Location/Qualifiers

FEATURES

source

1..300
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4324_C12_F24"
/issue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_id="wheat meiotic anther cDNA library"

/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; Plants were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary flowers. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared. cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5Kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 300;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||||
10 AGAACACACAGATTGCG 26

Db

RESULT 29
AM864697 304 bp mRNA linear EST 22-MAY-2000
LOCUS PM2-SN0018-160300-001-d01 SN0018 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AM864697
VERSION AM864697.1 GI:7998747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 304)

REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. U., Zago, N.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jorgensen, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

FEATURES
source 1 304
Location/Qualifiers
1 304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0018"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from CRSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 304;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||||
63 AGAACACACAGATTGCG 47

Db

RESULT 30
AV216565 312 bp mRNA linear EST 30-OCT-1999
LOCUS AV216565 RIKEN full-length enriched, ES cells Mus musculus cDNA
DEFINITION clone 2410164G16 3' similar to X57708 M. musculus RNA for pH 34, mRNA sequence.
ACCESSION AV216565
VERSION AV216565.1 GI:6157410
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 312)

REFERENCE
AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tomioka, N., Tomoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Yamamoto, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al. 1999)
Unpublished (1999)

CONTACT: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

EMAIL: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsuzaki, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kiteunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
Location/Qualifiers
1..312
1 312
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2410164G16"
/cell_type="ES cells"
/lab_host="SOLAR"
/clone_lib="RIKEN full-length enriched, ES cells"

FEATURES

source

/note="Site 1: XhoI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATTCTCGATTAAATTAATATCCCCCCCCCCC 3']".

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 312;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTGCG 18
Db 28 AAAACACACAGATTGCG 44

RESULT 31

AA787371 340 bp mRNA linear EST 31-JUL-1998
LOCUS n2g08a1.t1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone n2g08a1, mRNA sequence.

ACCESSION AA787371.1 GI:2847602
VERSION AA787371.1
KEYWORDS Emericella nidulans (anamorph: Aspergillus nidulans)
SOURCE Emericella nidulans
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eukaryotes; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 340)
AUTHORS Kudler, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broeou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou.edu

We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
High quality sequence stop: 326.

FEATURES

source

1..340
Location/Qualifiers
/organism="Emericella nidulans"
/mol_type="mRNA"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="n2g08a1"
/tissue_type="vegetative mycelia, asexual structures"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 340;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTGCG 17

Db 278 GAGAACACACAGATTGCG 294

RESULT 32

AM447131 351 bp mRNA linear EST 25-APR-2001
LOCUS AM447131
DEFINITION 88282 MARC 180V Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM447131
VERSION AM447131.1 GI:6988918
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 351)
Smith, T.P.L., Gross, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Caste, E., Wray, J.E., White, J., Cho, J., Fahrkrug, S.C., Bennett, G.L., Heaton, M.P., Jaeger, W.W., Rohrer, G.A., Chitko-McKown, C.G., Petree, G., Holt, I., Karaycheva, S., Liang, F., Quackenbush, J. and Keefe, D.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

TITLE

JOURNAL MEDLINE
PUBMED 11282978

COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACACGATGACCAT
BACKWARD: GTTTCACGACGACG
Plate: 63 row: L column: 5
Seq primer: ATTTCGTCGACCTATG.
Location/Qualifiers

FEATURES

source

1..351
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="PDH10B"
/clone_lib="MARC 180V"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 351;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACACACAGATTGCG 17
Db 74 GAGAACACACAGATTGCG 90

RESULT 33

AU257195 365 bp mRNA linear EST 25-APR-2002
LOCUS AU257195 3'-directed mouse cDNA library Mus musculus cDNA clone BEC0009976 3', mRNA sequence.

ACCESSION AU257195
VERSION AU257195.1 GI:20321577
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 365)
AUTHORS Kato, K. and Matoba, R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished (2002)
COMMENT Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatoc@is.tist-nara.ac.jp,
URL: http://love2.aist-nara.ac.jp/BEI/index.html.

FEATURES
source
1. .365
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="BBD0009976"
/tissue_type="brain"
/clone_lib="3-directed mouse cDNA library"

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 365;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTCCG 18
|||||
Db 306 AGACACACACAGATTCCG 322
|||||

RESULT 34
BZ651766/c 366 bp DNA linear GSS 29-JAN-2003
LOCUS OGANP82TC ZM_0.7_1.5_KB Zea mays genomic clone ZM5MA0102M20,
DEFINITION genomic survey sequence.
ACCESSION BZ651766
VERSION BZ651766.1 GI:28118555
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 366)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGANP82TM
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .366
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5MA0102M20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pGCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 365;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTCCG 18
|||||
Db 107 AGACACACACAGATTCCG 91
|||||

RESULT 35
BF851239 367 bp mRNA linear EST 16-JAN-2001
LOCUS IL5-EN0086-281100-291-g03 EN0086 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF851239
ACCESSION BF851239
VERSION BF851239.1 GI:12238401
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 367)
Dias, N.E., Garcia, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPER/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=IL5&2=IL5-EN0086-
281100-291-g03&t3=2000-11-28&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 336.
Location/Qualifiers
1. .367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0086"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESSES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 85.6%; Score 15.4; DB 2; Length 367;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTCCG 18
|||||
Db 218 AGACACACACAGATTCCG 234
|||||

RESULT 36
AY440562/c

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 877 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 389.
Location/Qualifiers
1..400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1625908"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: PT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL9W, testis NT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-726711, and 729096-73199. Subtraction by Benito
Soares and M. Fatima Ronaldo."

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 400;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAGATTCGC 18
|||
247 AGCACAACAGATTCGC 263
|||

RESULT 39 402 bp DNA linear GSS 20-OCT-2003
CG722358 119071F06.1BL Y1 1119 - Rescuem Grid AA Zee may's genomic
LOCUS survey sequence.
DEFINITION CG722358
VERSION CG722358.1 GI:37757127
KEYWORDS GSS.
SOURCE Zee may's
ORGANISM Zee may's
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zee.
1 (bases 1 to 402)
REFERENCE
1 Walbot, V.
Maize genomic sequences found using engineered Rescuem transposon
AUTHORS Unpublished (2001)
JOURNAL Contact: Walbot V
COMMENT Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 725 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 119071 row: 19
Class: transposon-tagged.
Location/Qualifiers
1..402
/organism="Zea mays"
/mol_type="genomic DNA"

FEATURES
source

/cultivar="mixed background W23/Al88/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="1119 - Rescuem Grid AA"
/note="Organ: leaf; Vector: Rescuem (engineered from
pBluescript backbone); Site 1: BamHI, Site 2: BglII;
Rescuem is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on Rescuem, go to the web
site 'www.zmnd.iastate.edu' and follow the links for
'Rescuem'. Grid AA was grown at UC San Diego in 2002. DNA
was extracted from leaf strips, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."

ORIGIN
Query Match 85.6%; Score 15.4; DB 9; Length 402;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACAACAGATTCG 17
|||
229 GAGAACAACAGATTCG 213
|||

Db 229 GAGAACAACAGATTCG 213
|||

RESULT 40 406 bp mRNA linear EST 29-JAN-1999
A1280419
LOCUS g195d08.x1 Soares NMHPu S1 Homo sapiens cDNA clone IMAGE:1880079
DEFINITION 3' similar to TR.F70193 F70193 INTEGRAL MEMBRANE GLYCOPROTEIN. ",
mRNA sequence.
VERSION A1280419.1 GI:3918652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 406)
REFERENCE
1 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 621 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 406.
Location/Qualifiers
1..406
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1880079"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NMHPu S1"
/note="Organ: mixed (see below); Vector: PT773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBH, pregnant uterus
NBH9U, and fetal heart NDHL9W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 406;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
DB 245 AGCACACACAGATTGCG 261

RESULT 41
A1220277 409 bp mRNA linear EST 30-NOV-1998
LOCUS
DEFINITION g973d03.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:1840805 3; similar to TR:P70193 P70193 INTEGRAL MEMBRANE
GLYCOPROTEIN.; mRNA sequence.

ACCESSION A1220277
VERSION A1220277.1 GI:3802480
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 409)
NCT-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lim.gov) for further information.
Insert Length: 1154 Std. Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 399.

FEATURES
source
1..409
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1840805"
/lab_host="DH10B"
/clone.lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL19W, testis NT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 287480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 409;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 170 AGCACACACAGATTGCG 186
RESULT 42
AA907628 413 bp mRNA linear EST 10-JUN-1998
LOCUS
DEFINITION cm10f08.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone

IMAGE:1540647 3; similar to TR:P70193 P70193 INTEGRAL MEMBRANE
GLYCOPROTEIN.; mRNA sequence.
AA907628
VERSION A1220277.1 GI:3043088
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 413)
NCT-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lim.gov) for further information.
Insert Length: 606 Std. Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 274.

FEATURES
source
1..413
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1540647"
/lab_host="DH10B"
/clone.lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL19W, testis NT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 287480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 413;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
DB 244 AGCACACACAGATTGCG 260

RESULT 43
BM173099 419 bp mRNA linear EST 04-DEC-2001
LOCUS
DEFINITION 900453 Avicennia marina leaf cDNA library Avicennia marina cDNA
clone Am900453 5; similar to unknown protein (AB025633) of
Arabidopsis thaliana, mRNA sequence.

ACCESSION BM173099
VERSION BM173099.1 GI:17312662
KEYWORDS EST.
SOURCE Avicennia marina
ORGANISM Avicennia marina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Acanthaceae; Acanthaceae incertae
sedis; Avicennia.
REFERENCE 1 (bases 1 to 419)
Parani, M., Mehra, P., Sivaprakash, K.R. and Parida, A.
Expressed sequence tags from the mangrove species Avicennia marina
Unpublished (2000)
JOURNAL Contact: Parani M / Parida A
COMMENT Department of Plant Molecular Biology

M. S. Swaminathan Research Foundation
111 Cross Street, Taramani Institutional Area, Chennai 600 113,
India
Tel: 91-44-2351319
Fax: 91-44-2351319
Email: mangrovegenes@msrf.res.in
BLAST search in BLASTX (Non-redundant) using default parameters as
on November 6th 2001
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES

source

```
1..419
/organism="Avicennia marina"
/mol_type="mRNA"
/strain="Pichavaram"
/db_xref="taxon:82927"
/clone="Am900453"
/sex="Hermaphrodite"
/tissue_type="leaf"
/dev_stage="1 year old plant"
/clone_lib="Avicennia marina leaf cDNA library"
/notes="A one-year-old plant from the natural mangrove
habitat in Pichavaram, Tamil Nadu, India was collected
and brought to the M.S. Swaminathan Research Foundation.
The whole plant was treated with the nutrient solution
supplemented with 500 mM NaCl for 48 hours. Poly(A+) mRNA
from the leaf tissue was purified and cDNA was prepared
using Superscript Lambda System (Life Technologies, Cat.
No. 19643-014). The cDNAs were size fractionated over
SizeSep-400 spun column (Amersham-Pharmacia Biotech Cat
No. 27-5105-01) and cloned in 5'Sal I - 3' Not I of
pSPORT1 (Life Technologies Cat. No. 15383-011). The ESTs
were sequenced from the 5' end using M13/pUC18 reverse
primer in an Automated Sequencer (ABI310, Applied
Biosystems) and submitted after editing to remove the
vector and adapter sequences. BLAST search in BLASTX
(Non-redundant) was carried out on 7th August, 2000 using
default parameters and the results reported under
putative identification/comment"
```

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 419;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTTCG 17
|||
82 GAGACACAGATTTCG 98

RESULT 44
AA812092 428 bp mRNA linear EST 19-FEB-1998
LOCUS AA812092
DEFINITION cb040a6.e1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:133810 3'
similar to TR:P70193 P70193 MEMBERANE GLYCOPROTEIN.; mRNA
sequence.
ACCESSION AA812092
VERSION AA812092.1 GI:2881703
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 428)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

JOURNAL

COMMENT

Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

FEATURES

source

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILN at:
www-bio.illn.gov/bdrp/image/image.html
Insert length: 1302 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham.
Location/Qualifiers

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 428;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTTCG 18
|||
213 AGCACACAGATTTCG 229

RESULT 45
BZ769778 430 bp DNA linear GSS 13-MAR-2003
LOCUS BZ769778
DEFINITION SALK_142711.47.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_142711.47.00.x, genomic
survey sequence.
ACCESSION BZ769778
VERSION BZ769778.1 GI:28943462
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 430)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shinn, P., Zimmerman, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers

TITLE

JOURNAL

COMMENT

Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

FEATURES

source

Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

```

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

```

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 430;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 17
 |||||
 DB 284 GAGAACACAGATTGCG 268

RESULT 46
 BP645308 431 bp mRNA linear EST 27-JUN-2004
 LOCUS BP645308 RAF1.9 Arabidopsis thaliana cDNA clone RAF1.9-68-F13 3',
 DEFINITION mRNA sequence.

ACCESSION BP645308 GI:49296778
 VERSION BP645308
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 431)

REFERENCE
 AUTHORS Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
 Nakaigima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,
 Hayashizaki, Y., Kawai, T., Carninci, P., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shingawa, A. and Shinzaki, K.
 Functional annotation of a full-length Arabidopsis cDNA collection
 Science 296 (5565), 141-145 (2002)
 MEDLINE 21932900
 PUBMED 11910074

COMMENT
 Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 reversed clone; please visit our web site
 (<http://pfigweb.gsc.riken.go.jp/>) for further details.

FEATURES

```

Location/Qualifiers
1..431
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF1.9-68-F13"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAF1.9"
/note="Site_1: BamHI, Site_2: SalI, Subtraction library"

```

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 431;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACAGATTGCG 18
 |||||
 DB 92 AGAACACAGATTGCG 108

RESULT 47
 BG543327/c 433 bp mRNA linear EST 01-MAY-2002
 LOCUS BG543327/c
 DEFINITION E0832 Chinese cabbage etiolated seedling library Brassica rapa
 subsp. pekinensis cDNA clone E0832, mRNA sequence.

ACCESSION BG543327 GI:20374307
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Brassica rapa subsp. pekinensis

REFERENCE
 AUTHORS Ryu, S.-H., Kang, U.-S., Kang, C.-h., Kim, C.-Y., Choi, Y.-J., Lee, S.-H.,
 Bank, J.-D., Lee, S.-Y., Cho, M.-J., and Lim, C.-O.
 Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA
 Unpublished (2001)
 CONTACT: Lim, C.O.
 Plant Molecular Biology & Biotechnology Research Centre
 Gyeongsang National University
 #900 Gazwa-dong, Jinju 660-701, Korea
 Tel: 82 55 751 6255
 Fax: 82 55 759 9363
 Email: collim@ongae.gsnu.ac.kr
 Seq primer: 17.

FEATURES

```

Location/Qualifiers
1..433
/organism="Brassica rapa subsp. pekinensis"
/mol_type="mRNA"
/cultivar="Jangwon"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="E0832"
/tissue_type="Etiolated seedling"
/lab_host="XL-1 Blue"
/clone_lib="Chinese cabbage etiolated seedling library"
/note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"

```

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 433;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACAGATTGCG 18
 |||||
 DB 159 AGAACACAGATTGCG 143

RESULT 48
 BM360109/c 437 bp mRNA linear EST 09-JAN-2002
 LOCUS GA_Ea0027G15+ Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Ea0027G15+, mRNA sequence.

ACCESSION BM360109 GI:18100855
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Gossypium arboreum

REFERENCE
 AUTHORS Henry, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
 Weng, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
 An integrated analysis of the genetics, development, and evolution
 of the cotton fiber
 Unpublished (2000)
 CONTACT: Wang RA
 Clemson University Genomics Institute
 Clemson University

100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total High Quality bases = 129
 Seq primer: TAATGCACTCAGCTATGAGG
 High quality sequence start: 15
 High quality sequence stop: 395.
 Location/Qualifiers
 1. 437
 /organism="Gossypium arboreum"
 /mol_type="mRNA"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone="GA_EA027G15r"
 /tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
 /lab_host="E. coli"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 437;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 |||||
 DB 39 AGAACACACAGATTGCG 23

RESULT 49 438 bp mRNA linear EST 30-SEP-1997
 LOCUS AA610098
 DEFINITION At19e07.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens CDNA clone IMAGE:1032132 3' similar to SW:SLIT_DROME P24014 SLIT PROTEIN PRCONSOR. /; mRNA sequence.
 ACCESSION AA610098
 VERSION AA610098.1 GI:2458526
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 438)
 Hillier, L., Allen, M., Bowles, J., Dubuque, T., Geisler, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheinberg, K., Sepcoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 399.
 Location/Qualifiers
 1. 438
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1032132"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"

FEATURES

source
 1. 438
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1032132"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares_total_fetus_Nb2HF8_9w"

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 438;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 |||||
 DB 244 AGAACACACAGATTGCG 260

RESULT 50 438 bp mRNA linear EST 26-MAY-2004
 LOCUS BU801478
 DEFINITION BU801478 unpublished oligo-capped CDNA library, stage L4
 Caenorhabditis elegans CDNA clone yk144912 5', mRNA sequence.
 ACCESSION BU801478
 VERSION BU801478.1 GI:47705145
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.
 1 (bases 1 to 438)
 Kohara, Y., Shin, J.T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y., and Sugano, S.
 A complementary view of the C. elegans genome
 Unpublished (2002)
 Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers
 1. 438
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk144912"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L4"
 /clone_lib="unpublished oligo-capped CDNA library, stage L4"

FEATURES

source
 1. 438
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk144912"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L4"
 /clone_lib="unpublished oligo-capped CDNA library, stage L4"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 438;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 |||||
 DB 142 AGAACACACAGATTGCG 158

Search completed: December 3, 2004, 05:47:11
 Job time: 2354.45 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: December 2, 2004, 20:35:25 ; Search time 257.211 Seconds
(without alignments)
367.363 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18

Sequence: 1 gagaaacaacagatcgc 18

Scoring table:

IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

N_Geneseq_238604:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	16.4	91.1	65589	6	ABA90521.23	Continuation (24 o
2	16	88.9	19	6	ABN84784	Abn84784 Primer us
3	16	88.9	66479	6	ABQ80567	Abq80567 Primer us
4	16	88.9	66479	6	ABQ80566	Abq80566 Mutant hu
5	16	88.9	66479	6	ABQ80568	Abq80568 Mutant hu
6	16	88.9	66479	6	ABQ80565	Abq80565 Mutant hu
7	15.4	88.6	485	6	ABL37566	Ab137566 Human col
8	15.4	88.6	513	8	ACA32902	ACA32902 Prokaryot
9	15.4	88.6	534	10	ADH84638	Adh84638 Enterococ
C 10	15.4	88.6	1470	4	AA534481	Aa534481 E. coli D
11	15.4	88.6	1822	10	ADH72695	Adh72695 Human en
C 12	15.4	88.6	1944	5	AA588539	Aa588539 DNA encod
C 13	15.4	88.6	1944	5	AA590079	Aa590079 DNA encod
C 14	15.4	88.6	1944	5	AA593115	Aa593115 DNA encod
C 15	15.4	88.6	2186	4	AAH48474	Aah48474 Escherich
16	15.4	85.6	2620	10	ADH72674	Adh72674 Human en
17	15.4	85.6	2747	10	ADH72675	Adh72675 Human en
18	15.4	85.6	2789	10	ADH72693	Adh72693 Human en
19	15.4	85.6	3178	2	AAV30824	Aav30824 Mitochond
20	15.4	85.6	3178	2	AAV42018	Aav42018 Glycerol-
21	15.4	85.6	3178	2	AAV35739	Aav35739 Klebsiell

22	15.4	85.6	3178	2	AA80613	CUT2 gene
23	15.4	85.6	3333	12	ADH72187	Human gen
24	15.4	85.6	3333	12	ADH72193	Human gen
25	15.4	85.6	3368	10	ADH72673	Human end
26	15.4	85.6	3470	10	ADH72686	Human end
27	15.4	85.6	3741	10	ADH72691	Human end
28	15.4	85.6	3858	10	ADH72694	Human end
29	15.4	85.6	3886	10	ADH72694	Human end
30	15.4	85.6	4067	10	ADH72687	Human end
31	15.4	85.6	4331	10	ADH72692	Human end
32	15.4	85.6	4404	10	ADH72671	Human end
33	15.4	85.6	4717	6	ABK99971	DNA encod
34	15.4	85.6	4751	10	ADH72672	Human end
35	15.4	85.6	4759	6	ABK99957	DNA encod
36	15.4	85.6	4762	8	AA157275	LIG-1/hu
37	15.4	85.6	4762	12	ADH72191	Human gen
38	15.4	85.6	5101	6	ABV99394	Human NOV
39	15.4	85.6	5101	12	ADH72189	Human gen
40	15.4	85.6	5188	10	ADH72682	Human end
41	15.4	85.6	5188	10	ADH72688	Human end
42	15.4	85.6	5263	10	ADH72689	Human end
43	15.4	85.6	5273	10	ADH72689	Human end
44	15.4	85.6	5287	12	ADH05928	Antipori
45	15.4	85.6	5332	10	ADH72676	Human end
46	15.4	85.6	5337	10	ADH72666	Human end
47	15.4	85.6	5353	10	ADH72667	Human end
48	15.4	85.6	5639	10	ADH72681	Human end
49	15.4	85.6	5684	10	ADH72685	Human end
50	15.4	85.6	5857	10	ADH72670	Human end
51	15.4	85.6	5921	10	ADH72684	Human end
52	15.4	85.6	5922	10	ADH72683	Human end
53	15.4	85.6	7043	6	AAK13225	Enterococ
54	15.4	85.6	7043	6	ABK99020	Enterococ
55	15.4	85.6	7462	2	AAQ35153	Jaagsiek
56	15.4	85.6	204	10	ADH181939	Cell divi
57	15.4	85.6	485	3	AAK62754	Cell divi
58	15.4	85.6	485	6	ABK17019	Pinus rad
59	15.4	85.6	485	10	ADH75420	Pinus rad
60	15.4	85.6	737	6	ABQ65500	Arabidops
61	15.4	85.6	991	3	AAK34103	Arabidops
62	15.4	85.6	1455	6	ABK21846	Candida a
63	15.4	85.6	95769	8	ADH68659	Arabidops
64	14.8	82.2	352	6	ABQ98993	Human CRP
65	14.8	82.2	372	10	ADH87115	Human pan
66	14.8	82.2	372	10	ADH87114	Human pan
67	14.8	82.2	399	10	ADH86036	Aspergill
68	14.8	82.2	438	8	ACA45733	Prokaryot
69	14.8	82.2	465	9	ACH14633	Human acu
70	14.8	82.2	467	10	ADC03438	Banana fl
71	14.8	82.2	491	12	ADH43785	Plant CDN
72	14.8	82.2	491	8	ABO83836	Aspergill
73	14.8	82.2	544	11	ABV77520	A. cryzae
74	14.8	82.2	544	11	ADH45267	Insect re
75	14.8	82.2	602	4	ABH12223	Drosophil
76	14.8	82.2	634	6	ABK17189	Human dlt
77	14.8	82.2	676	6	ACA04567	DNA enco
78	14.8	82.2	795	3	AAK12329	Aspergill
79	14.8	82.2	858	3	ACA53904	Prokaryot
80	14.8	82.2	943	6	ABQ38422	Oligonuc
81	14.8	82.2	943	6	ABQ38423	Oligonuc
82	14.8	82.2	1014	3	AAK08052	Fusarium
83	14.8	82.2	1038	3	AAK63736	Arabidops
84	14.8	82.2	1125	3	AAK43034	Arabidops
85	14.8	82.2	1193	4	AAK07510	Arabidops
86	14.8	82.2	1212	12	ADH39796	Plant CDN
87	14.8	82.2	1359	4	AAK60735	Pseudomona
88	14.8	82.2	1360	10	ADH20193	Cat pregn
89	14.8	82.2	1360	10	ADH1604	Peline pr
90	14.8	82.2	1389	8	ACA53957	Prokaryot
91	14.8	82.2	1478	3	AAK46267	Arabidops
92	14.8	82.2	1480	3	AAK35460	Arabidops
93	14.8	82.2	1964	10	ADH48694	Human gen
94	14.8	82.2	1964	10	ADH48694	Human gen

C 95	14.8	82.2	2000	8	ADA73070	AdA73070	Rice gene	159	14.4	80.0	1106	3	AAC47255	Aac47255	Arabidops
C 96	14.8	82.2	2157	4	ABE12755	ABe12755	Drosophi	159	14.4	80.0	1106	3	AAC38805	Aac38805	Arabidops
C 97	14.8	82.2	2458	8	ACA8451	ACa8451	Prokaryot	159	14.4	80.0	1109	3	AAC38856	Aac38856	Arabidops
C 98	14.8	82.2	2529	12	ADN72544	ADN72544	Thale cre	171	14.4	80.0	1159	8	AAa87700	AaA87700	Human sec
C 99	14.8	82.2	2602	4	AB112222	AB112222	Drosophi	171	14.4	80.0	1159	8	ADA40321	AdA40321	Human sec
C 100	14.8	82.2	2712	8	ACA21040	ACa21040	Prokaryot	173	14.4	80.0	1159	10	ADA56548	AdA56548	Gene enco
C 101	14.8	82.2	2727	9	ADCA11755	ADCA11755	DNA enco	173	14.4	80.0	1200	3	AAa87678	AaA87678	Human sec
C 102	14.8	82.2	2831	4	ADDS5832	ADDS5832	Arabidops	175	14.4	80.0	1200	8	ADA39768	AdA39768	Human sec
C 103	14.8	82.2	2831	10	ADDS5693	ADDS5693	Thalecres	176	14.4	80.0	1200	10	ADA55959	AdA55959	Gene enco
C 104	14.8	82.2	2831	10	ADDD3000	ADDD3000	Plant yle	C 177	14.4	80.0	1357	5	AAS69904	AaS69904	DNA enco
C 105	14.8	82.2	2831	12	ADT41614	ADt41614	Plant tra	178	14.4	80.0	1433	11	ADM07034	AdM07034	Aspergill
C 106	14.8	82.2	2831	12	ADDO03562	ADDO03562	Thalecres	179	14.4	80.0	1433	11	ADM07093	AdM07093	Aspergill
C 107	14.8	82.2	2831	12	ADDO01870	ADDO01870	Thalecres	180	14.4	80.0	1423	4	ABL17355	ABl17355	Human mus
C 108	14.8	82.2	2831	12	ADDO01630	ADDO01630	Transcrip	181	14.4	80.0	1423	8	ABX60383	ABx60383	DNA enco
C 109	14.8	82.2	3081	4	AB102575	AB102575	Drosophi	C 182	14.4	80.0	1423	12	ADBJ31133	ADbJ31133	Human mus
C 110	14.8	82.2	3081	10	ADK11353	ADk11353	Drosophi	C 183	14.4	80.0	1462	2	AAV75082	AAv75082	Staphyloc
C 111	14.8	82.2	3098	4	AB113617	AB113617	Drosophi	C 184	14.4	80.0	1722	6	ABE15789	ABe15789	Magnaport
C 112	14.8	82.2	3511	4	AB111009	AB111009	Drosophi	C 185	14.4	80.0	1830	4	AAI60414	AAI60414	Human pol
C 113	14.8	82.2	3511	10	ADK11438	ADk11438	Drosophi	C 186	14.4	80.0	1833	4	AAI58638	AAI58638	Human pol
C 114	14.8	82.2	3521	4	AB112846	AB112846	Drosophi	187	14.4	80.0	1833	5	ADQ98846	ADq98846	DNA enco
C 115	14.8	82.2	3593	4	AB120918	AB120918	Drosophi	188	14.4	80.0	1833	5	ADB48606	ADb48606	Novel hum
C 116	14.8	82.2	4197	4	AA807404	AA807404	Arabidops	C 189	14.4	80.0	1878	8	ADA68528	ADa68528	Arabidops
C 117	14.8	82.2	4197	4	AA807406	AA807406	Arabidops	C 190	14.4	80.0	1950	2	AAK90402	AAK90402	Nicotiana
C 118	14.8	82.2	4197	4	AA807571	AA807571	Arabidops	C 191	14.4	80.0	2137	12	ADN88789	ADn88789	Nicotiana
C 119	14.8	82.2	4197	4	AA807405	AA807405	Arabidops	C 192	14.4	80.0	2271	4	AAH31358	AAh31358	Human sec
C 120	14.8	82.2	4197	4	AA807403	AA807403	Arabidops	193	14.4	80.0	2359	10	ADB622773	ADb622773	Human cDN
C 121	14.8	82.2	4197	4	AA807402	AA807402	Arabidops	194	14.4	80.0	2390	4	ADD08046	ADD08046	Human ext
C 122	14.8	82.2	4624	4	ABE122501	AB122501	Drosophi	195	14.4	80.0	2433	11	ADMC07092	ADmC07092	Aspergill
C 123	14.8	82.2	4694	8	ACA45314	ACa45314	Prokaryot	196	14.4	80.0	2553	6	AAI70607	AAI70607	Pichia ec
C 124	14.8	82.2	5531	4	ABE117158	AB117158	Drosophi	C 197	14.4	80.0	2958	3	AAC48752	ACa48752	Arabidops
C 125	14.8	82.2	5844	4	ABE11008	AB111008	Drosophi	198	14.4	80.0	3078	8	ACC69470	ACc69470	Human mal
C 126	14.8	82.2	6636	6	ABEN80023	ABn80023	Human che	199	14.4	80.0	3125	6	ABO54968	ABo54968	Human ova
C 127	14.8	82.2	7146	4	AB1131616	AB113166	Drosophi	200	14.4	80.0	3266	2	AAK52267	AAK52267	Protein P
C 128	14.8	82.2	7747	4	AA807401	AA807401	Arabidops	201	14.4	80.0	3266	3	AAC58590	AAc58590	Human PRO
C 129	14.8	82.2	8043	4	AB1025514	AB1025514	Drosophi	202	14.4	80.0	3266	3	ADCT78621	ADcT78621	Human PRO
C 130	14.8	82.2	8202	4	ABE019180	AB1019180	Drosophi	203	14.4	80.0	3266	4	AAAF72425	AAAF72425	Human PRO
C 131	14.8	82.2	8731	4	ABE111358	AB111358	Drosophi	204	14.4	80.0	3266	4	AAE211428	AAe211428	Human cDN
C 132	14.8	82.2	35724	4	AAC85022	Aac85022	Adenoviru	205	14.4	80.0	3266	8	ACA60256	ACa60256	Human cDN
C 133	14.8	82.2	36114	4	AAC85025	Aac85025	Adenoviru	206	14.4	80.0	3266	8	ACD07657	ACd07657	Human hum
C 134	14.8	82.2	44377	2	AAAT80414	AAe80414	Platenoli	207	14.4	80.0	3266	8	ACA03787	ACa03787	DNA enco
C 135	14.8	82.2	44377	2	AAT78508	AAe78508	Platenoli	208	14.4	80.0	3266	8	ABX711705	ABx711705	Human cDN
C 136	14.8	82.2	84428	12	ADMA45913	ADm45913	Streptomy	209	14.4	80.0	3266	8	ACH07036	ACH07036	Human sec
C 137	14.8	82.2	110000	2	AAAX20448	AAAX20448	Continuation (6 of	210	14.4	80.0	3266	8	ABX89325	ABX89325	DNA enco
C 138	14.8	82.2	110000	4	AAI95682_75	AAI95682_75	Continuation (36 o	211	14.4	80.0	3266	8	ACDA11979	ACdA11979	Human sec
C 139	14.8	82.2	110000	4	AAI95683_35	AAI95683_35	Continuation (36 o	212	14.4	80.0	3266	8	ABX66273	ABx66273	Human sec
C 140	14.8	82.2	111309	2	AAAX20250	AAAX20250	Borrella	213	14.4	80.0	3266	8	ACA05554	ACa05554	DNA enco
C 141	14.4	80.0	60	9	AAD58072	AAe58072	ErBb2 bin	214	14.4	80.0	3266	8	ACA04208	ACa04208	Human cDN
C 142	14.4	80.0	174	12	ACH86795	ACH86795	Human gen	215	14.4	80.0	3266	8	ACA55065	ACa55065	Novel hum
C 143	14.4	80.0	174	12	ADO26778	ADO26778	Human INS	216	14.4	80.0	3266	9	ADA45888	ADa45888	Novel hum
C 144	14.4	80.0	369	4	AAH69054	AAH69054	Human cer	217	14.4	80.0	3266	9	ADA76319	ADa76319	Human PRO
C 145	14.4	80.0	439	4	AAH71317	AAH71317	Human cer	218	14.4	80.0	3266	9	ADAB29514	ADaB29514	Human sec
C 146	14.4	80.0	439	10	ADH74705	ADh74705	Human can	219	14.4	80.0	3266	9	ADA18969	ADa18969	Human PRO
C 147	14.4	80.0	441	4	AA185921	AA185921	Human pol	220	14.4	80.0	3266	9	ADA61552	ADa61552	Novel hum
C 148	14.4	80.0	478	8	ACH15202	ACH15202	Human adu	221	14.4	80.0	3266	9	ADB11937	ADb11937	Novel hum
C 149	14.4	80.0	493	9	ACH44748	ACH44748	Human foe	222	14.4	80.0	3266	9	ADB27918	ADb27918	DNA enco
C 150	14.4	80.0	510	4	AAH70717	AAH70717	Human cer	223	14.4	80.0	3266	9	ADBA6337	ADBA6337	Novel hum
C 151	14.4	80.0	530	12	ACH73064	ACH73064	Human gen	224	14.4	80.0	3266	9	ADBA15961	ADBA15961	Human PRO
C 152	14.4	80.0	553	9	ACH24266	ACH24266	Human adu	225	14.4	80.0	3266	9	ADDA7747	ADDA7747	Human PRO
C 153	14.4	80.0	591	8	ABZ75080	ABZ75080	Mouse emb	226	14.4	80.0	3266	9	ADA18371	ADA18371	Human sec
C 154	14.4	80.0	591	8	ABZ75088	ABZ75088	Mouse emb	227	14.4	80.0	3266	9	ACDE7046	ACDe7046	Human cDN
C 155	14.4	80.0	591	11	ABD04328	ABd04328	Pseudomon	228	14.4	80.0	3266	9	ADAE67542	ADAE67542	Human PRO
C 156	14.4	80.0	648	8	ACF72703	ACf72703	Staphyloc	229	14.4	80.0	3266	9	ADB930549	ADb930549	DNA enco
C 157	14.4	80.0	684	6	ABZ13158	ABz13158	Arabidops	230	14.4	80.0	3266	9	ADAE5845	ADAE5845	Novel hum
C 158	14.4	80.0	684	6	ADG87922	ADg87922	A. thalia	231	14.4	80.0	3266	9	ADAE87057	ADAE87057	Human PRO
C 159	14.4	80.0	684	8	ADA68387	ADA68387	Arabidops	232	14.4	80.0	3266	9	ADAA79361	ADAA79361	Human PRO
C 160	14.4	80.0	684	11	ABD04239	ABD04239	Pseudomon	233	14.4	80.0	3266	9	ADAE87500	ADAE87500	Novel hum
C 161	14.4	80.0	696	3	AAAC39611	AAc39611	Arabidops	234	14.4	80.0	3266	9	ADBI6702	ADBI6702	Human PRO
C 162	14.4	80.0	701	3	AAAC50385	AAc50385	Arabidops	235	14.4	80.0	3266	9	ACDB3207	ACDB3207	Human PRO
C 163	14.4	80.0	740	3	AAAO1696	AAa01696	Human col	236	14.4	80.0	3266	9	ADAI61346	ADAI61346	Human sec
C 164	14.4	80.0	808	11	ADM83440	ADM83440	Mycobacte	237	14.4	80.0	3266	9	ADA91794	ADA91794	Novel hum
C 165	14.4	80.0	808	8	AAH72679	AAh72679	Human cer	238	14.4	80.0	3266	9	ADBA14657	ADBA14657	Human PRO
C 166	14.4	80.0	819	8	ACF73930	ACf73930	Staphyloc	239	14.4	80.0	3266	9	ADBI18818	ADBI18818	Novel hum
C 167	14.4	80.0	867	11	ABD04519	ABD04519	Pseudomon	240	14.4	80.0	3266	9	ADA94033	ADa94033	Human PRO

241	14.4	80.0	3296	9	ADB19929	Novel hum	314	14.4	80.0	3296	10	ADC40927	Human sec
242	14.4	80.0	3296	9	ADB13241	Human PRO	315	14.4	80.0	3296	10	ADC19584	Human sec
243	14.4	80.0	3296	9	ACD98608	Novel hum	316	14.4	80.0	3296	10	ADC34032	Human sec
244	14.4	80.0	3296	9	ADA74495	Human PRO	317	14.4	80.0	3296	10	ADC13102	Human sec
245	14.4	80.0	3296	9	ADA42491	Human sec	318	14.4	80.0	3296	10	ADC50410	Novel hum
246	14.4	80.0	3296	9	ADB24728	Human PRO	319	14.4	80.0	3296	10	ADC71957	Novel hum
247	14.4	80.0	3296	9	ADA82252	Human PRO	320	14.4	80.0	3296	10	ADC59936	Novel hum
248	14.4	80.0	3296	9	ADA75215	Human PRO	321	14.4	80.0	3296	10	ADC52943	Novel hum
249	14.4	80.0	3296	9	ADA85293	Novel hum	322	14.4	80.0	3296	10	ADC57297	Novel hum
250	14.4	80.0	3296	9	ADA84741	Novel hum	323	14.4	80.0	3296	10	ADC60488	Novel hum
251	14.4	80.0	3296	9	ACD23385	Human PRO	324	14.4	80.0	3296	10	ADC50963	Novel hum
252	14.4	80.0	3296	9	ADB23997	CDNA enco	325	14.4	80.0	3296	10	ADC65490	Novel hum
253	14.4	80.0	3296	9	ADA80525	Human PRO	326	14.4	80.0	3296	10	ADC54588	Novel hum
254	14.4	80.0	3296	9	ADA75767	Human PRO	327	14.4	80.0	3296	10	ADC53549	Novel hum
255	14.4	80.0	3296	9	ADA46992	Human PRO	328	14.4	80.0	3296	10	ADC59072	Novel hum
256	14.4	80.0	3296	9	ADB25288	Human PRO	329	14.4	80.0	3296	10	ADC55950	Novel hum
257	14.4	80.0	3296	9	ADA93464	Human PRO	330	14.4	80.0	3296	10	ADC58520	Novel hum
258	14.4	80.0	3296	9	ADB26814	CDNA enco	331	14.4	80.0	3296	10	ADC12554	Novel hum
259	14.4	80.0	3296	9	ADB31101	CDNA enco	332	14.4	80.0	3296	10	ADC03194	Novel hum
260	14.4	80.0	3296	9	ADA61029	Homo sapi	333	14.4	80.0	3296	10	ADC90186	Novel hum
261	14.4	80.0	3296	9	ADB24176	Human PRO	334	14.4	80.0	3296	10	ADC69605	Novel hum
262	14.4	80.0	3296	9	ADA96505	Human PRO	335	14.4	80.0	3296	10	ADC48494	Novel hum
263	14.4	80.0	3296	9	ADA81077	Human PRO	336	14.4	80.0	3296	10	ADD10023	Novel hum
264	14.4	80.0	3296	9	ADA95953	Human PRO	337	14.4	80.0	3296	10	ADD04598	Novel hum
265	14.4	80.0	3296	9	ADB26262	CDNA enco	338	14.4	80.0	3296	10	ADC80554	Novel hum
266	14.4	80.0	3296	9	ADB21747	Novel hum	339	14.4	80.0	3296	10	ADD11061	Novel hum
267	14.4	80.0	3296	9	ADA77526	Human PRO	340	14.4	80.0	3296	10	ADC47942	Novel hum
268	14.4	80.0	3296	9	ADB18266	CDNA enco	341	14.4	80.0	3296	10	ADD05109	Novel hum
269	14.4	80.0	3296	9	ADA86949	Novel hum	342	14.4	80.0	3296	10	ADC80002	Novel hum
270	14.4	80.0	3296	9	ADA16770	Human sec	343	14.4	80.0	3296	10	ADD09471	Novel hum
271	14.4	80.0	3296	9	ADA13199	Human sec	344	14.4	80.0	3296	10	ADD04115	Novel hum
272	14.4	80.0	3296	9	ADA42067	Human sec	345	14.4	80.0	3296	10	ADD03691	Novel hum
273	14.4	80.0	3296	9	ADA88052	Novel hum	346	14.4	80.0	3296	10	ADD41184	Novel hum
274	14.4	80.0	3296	9	ADA46440	Novel hum	347	14.4	80.0	3296	10	ADD52323	Novel hum
275	14.4	80.0	3296	9	ADA17414	Human sec	348	14.4	80.0	3296	10	ADD53063	Novel hum
276	14.4	80.0	3296	9	ADA42817	Human sec	349	14.4	80.0	3296	10	ADD53615	Novel hum
277	14.4	80.0	3296	9	ADB28470	CDNA enco	350	14.4	80.0	3296	10	ADD51771	Novel hum
278	14.4	80.0	3296	9	ADB29022	CDNA enco	351	14.4	80.0	3296	10	ADD02570	Novel hum
279	14.4	80.0	3296	9	ADA76974	Human PRO	352	14.4	80.0	3296	10	ADD02004	Novel hum
280	14.4	80.0	3296	9	ADA88604	Novel hum	353	14.4	80.0	3296	10	ADD54186	Novel hum
281	14.4	80.0	3296	9	ADA97609	Human PRO	354	14.4	80.0	3296	10	ADD92503	Novel hum
282	14.4	80.0	3296	9	ADB27366	CDNA enco	355	14.4	80.0	3296	10	ADD91399	Novel hum
283	14.4	80.0	3296	9	ADB22299	Novel hum	356	14.4	80.0	3296	10	ADB04013	Novel hum
284	14.4	80.0	3296	9	ACD23747	Human PRO	357	14.4	80.0	3296	10	ADB32310	Novel hum
285	14.4	80.0	3296	9	ADA66990	Human PRO	358	14.4	80.0	3296	10	ADB22242	Novel hum
286	14.4	80.0	3296	9	ADB22851	Human PRO	359	14.4	80.0	3296	10	ADD79466	CDNA enco
287	14.4	80.0	3296	9	ADB23624	Human PRO	360	14.4	80.0	3296	10	ADB42002	Novel hum
288	14.4	80.0	3296	9	ADA92246	Novel hum	361	14.4	80.0	3296	10	ADB17819	Novel hum
289	14.4	80.0	3296	9	ADB15409	Human PRO	362	14.4	80.0	3296	10	ADB31951	Novel hum
290	14.4	80.0	3296	9	ADB38661	Novel hum	363	14.4	80.0	3296	10	ADB33414	Novel hum
291	14.4	80.0	3296	9	ADB38109	Novel hum	364	14.4	80.0	3296	10	ADB33966	Novel hum
292	14.4	80.0	3296	9	ADB66581	Novel hum	365	14.4	80.0	3296	10	ADB33966	Novel hum
293	14.4	80.0	3296	10	ADB89661	Human PRO	366	14.4	80.0	3296	10	ADD93055	Novel hum
294	14.4	80.0	3296	10	ADB90393	Human PRO	367	14.4	80.0	3296	10	ADD93055	Novel hum
295	14.4	80.0	3296	10	ADB7835	Human sec	368	14.4	80.0	3296	10	ADB19475	Novel hum
296	14.4	80.0	3296	10	ADB39494	Novel hum	369	14.4	80.0	3296	10	ADB34943	Novel hum
297	14.4	80.0	3296	10	ADB74971	Human sec	370	14.4	80.0	3296	10	ADB18923	Novel hum
298	14.4	80.0	3296	10	ADB47117	Novel hum	371	14.4	80.0	3296	10	ADB43119	Novel hum
299	14.4	80.0	3296	10	ADB86724	Human PRO	372	14.4	80.0	3296	10	ADD22794	Novel hum
300	14.4	80.0	3296	10	ADB77329	Novel hum	373	14.4	80.0	3296	10	ADB278912	Novel hum
301	14.4	80.0	3296	10	ADB33486	Human PRO	374	14.4	80.0	3296	10	ADB32862	Novel hum
302	14.4	80.0	3296	10	ADB35590	Human PRO	375	14.4	80.0	3296	10	ADB42554	Novel hum
303	14.4	80.0	3296	10	ADB35934	Human PRO	376	14.4	80.0	3296	10	ADB42554	Novel hum
304	14.4	80.0	3296	10	ADB35038	Human PRO	377	14.4	80.0	3296	10	ADB80570	Novel hum
305	14.4	80.0	3296	10	ADB36142	Human PRO	378	14.4	80.0	3296	10	ADB89598	Novel hum
306	14.4	80.0	3296	10	ADB46537	Novel hum	379	14.4	80.0	3296	10	ADB40882	Novel hum
307	14.4	80.0	3296	10	ADB28618	Human sec	380	14.4	80.0	3296	10	ADB4681	Novel hum
308	14.4	80.0	3296	10	ADC39818	Human sec	381	14.4	80.0	3296	10	ADB92810	Novel hum
309	14.4	80.0	3296	10	ADC40332	Human sec	382	14.4	80.0	3296	10	ADB21519	Novel hum
310	14.4	80.0	3296	10	ADC19156	Human sec	383	14.4	80.0	3296	10	ADC23160	Novel hum
311	14.4	80.0	3296	10	ADC34456	Human sec	384	14.4	80.0	3296	10	ADP97495	Novel hum
312	14.4	80.0	3296	10	ADC29511	Human sec	385	14.4	80.0	3296	10	ADB80559	Novel hum
313	14.4	80.0	3296	10	ADC29042	Human sec	386	14.4	80.0	3296	10	ADB80007	Novel hum
												ADB59426	Human sec

387	14.4	80.0	3296	10	ADH55299	Novel hum	460	14.4	80.0	3296	12	ADG82215	Novel PRO
388	14.4	80.0	3296	10	ADH55851	Novel hum	461	14.4	80.0	3296	12	ADG57454	Novel hum
389	14.4	80.0	3296	10	ADH38205	Novel hum	462	14.4	80.0	3296	12	ADG56902	Novel hum
390	14.4	80.0	3296	10	ADH64070	Novel hum	463	14.4	80.0	3296	12	ADG55798	Novel hum
391	14.4	80.0	3296	10	ADH65019	Novel hum	464	14.4	80.0	3296	12	ADG58558	Novel hum
392	14.4	80.0	3296	10	ADH63518	Novel hum	465	14.4	80.0	3296	12	ADG70924	Novel hum
393	14.4	80.0	3296	10	ADH81932	Novel hum	466	14.4	80.0	3296	12	ADG92833	Novel hum
394	14.4	80.0	3296	10	ADH81380	Novel hum	467	14.4	80.0	3296	12	ADG58006	Novel hum
395	14.4	80.0	3296	10	ADH81592	Novel hum	468	14.4	80.0	3296	12	ADG53590	Novel hum
396	14.4	80.0	3296	10	ADH24037	Novel hum	469	14.4	80.0	3296	12	ADG71476	Novel hum
397	14.4	80.0	3296	10	ADH58549	Novel hum	470	14.4	80.0	3296	12	ADG81663	Novel hum
398	14.4	80.0	3296	10	ADH64178	Novel hum	471	14.4	80.0	3296	12	ADH30625	Novel hum
399	14.4	80.0	3296	10	ADH26473	Novel hum	472	14.4	80.0	3296	12	ADH11992	Novel hum
400	14.4	80.0	3296	11	ADH62549	Novel hum	473	14.4	80.0	3296	12	ADG52414	Novel hum
401	14.4	80.0	3296	11	ADH59448	Novel hum	474	14.4	80.0	3296	12	ADG54142	Novel hum
402	14.4	80.0	3296	11	ADH16577	Novel hum	475	14.4	80.0	3296	12	ADG81111	Novel hum
403	14.4	80.0	3296	11	ADH15396	Novel hum	476	14.4	80.0	3296	12	ADG56350	Novel hum
404	14.4	80.0	3296	11	ADH14844	Novel hum	477	14.4	80.0	3296	12	ADH12616	Novel hum
405	14.4	80.0	3296	11	ADH61106	Novel hum	478	14.4	80.0	3296	12	ADG61462	Novel hum
406	14.4	80.0	3296	12	ADH79388	Novel hum	479	14.4	80.0	3296	12	ADH28549	Novel hum
407	14.4	80.0	3296	12	ADH76554	Novel hum	480	14.4	80.0	3296	12	ADG54694	Novel hum
408	14.4	80.0	3296	12	ADH87918	Novel hum	481	14.4	80.0	3296	12	ADG59734	Novel hum
409	14.4	80.0	3296	12	ADH63222	Novel hum	482	14.4	80.0	3296	12	ADH20622	Novel hum
410	14.4	80.0	3296	12	ADH79812	Novel hum	483	14.4	80.0	3296	12	ADH07477	Novel hum
411	14.4	80.0	3296	12	ADH75770	Novel hum	484	14.4	80.0	3296	12	ADH60022	Novel hum
412	14.4	80.0	3296	12	ADH73488	Novel hum	485	14.4	80.0	3296	12	ADH07050	Novel hum
413	14.4	80.0	3296	12	ADH23346	Novel hum	486	14.4	80.0	3296	12	ADH18158	Novel hum
414	14.4	80.0	3296	12	ADH23898	Novel hum	487	14.4	80.0	3296	12	ADH18792	Novel hum
415	14.4	80.0	3296	12	ADH24541	Novel hum	488	14.4	80.0	3296	12	ADH16512	Novel hum
416	14.4	80.0	3296	12	ADH87366	Novel hum	489	14.4	80.0	3296	12	ADH13771	Novel hum
417	14.4	80.0	3296	12	ADH89232	Novel hum	490	14.4	80.0	3296	12	ADH09901	Novel hum
418	14.4	80.0	3296	12	ADH74023	Novel hum	491	14.4	80.0	3296	12	ADH97571	Novel hum
419	14.4	80.0	3296	12	ADH18371	Novel hum	492	14.4	80.0	3296	12	ADH15372	Novel hum
420	14.4	80.0	3296	12	ADH88680	Novel hum	493	14.4	80.0	3296	12	ADH08249	Novel hum
421	14.4	80.0	3296	12	ADH95977	Novel hum	494	14.4	80.0	3296	12	ADH165939	Novel hum
422	14.4	80.0	3296	12	ADH84700	Novel hum	495	14.4	80.0	3296	12	ADH14704	Novel hum
423	14.4	80.0	3296	12	ADH91111	Novel hum	496	14.4	80.0	3296	12	ADH60682	Novel hum
424	14.4	80.0	3296	12	ADH93252	Novel hum	497	14.4	80.0	3296	12	ADH18299	Novel hum
425	14.4	80.0	3296	12	ADH93362	Novel hum	498	14.4	80.0	3296	12	ADH19739	Novel hum
426	14.4	80.0	3296	12	ADH34943	Novel hum	499	14.4	80.0	3296	12	ADH08932	Novel hum
427	14.4	80.0	3296	12	ADH38866	Novel hum	500	14.4	80.0	3296	12	ADH25273	Novel hum
428	14.4	80.0	3296	12	ADH32258	Novel hum	501	14.4	80.0	3296	12	ADH63580	Novel hum
429	14.4	80.0	3296	12	ADH90559	Novel hum	502	14.4	80.0	3296	12	ADH30023	Novel hum
430	14.4	80.0	3296	12	ADH91706	Novel hum	503	14.4	80.0	3296	12	ADH77475	Novel hum
431	14.4	80.0	3296	12	ADH99123	Novel hum	504	14.4	80.0	3296	12	ADH65597	Novel hum
432	14.4	80.0	3296	12	ADH40593	Novel hum	505	14.4	80.0	3296	12	ADH27733	Novel hum
433	14.4	80.0	3296	12	ADH73987	Novel hum	506	14.4	80.0	3296	12	ADH42457	Novel hum
434	14.4	80.0	3296	12	ADH02285	Novel hum	507	14.4	80.0	3296	12	ADH06345	Novel hum
435	14.4	80.0	3296	12	ADH22071	Novel hum	508	14.4	80.0	3296	12	ADH28319	Novel hum
436	14.4	80.0	3296	12	ADH20141	Novel hum	509	14.4	80.0	3296	12	ADH63903	Novel hum
437	14.4	80.0	3296	12	ADH98047	Novel hum	510	14.4	80.0	3296	12	ADH69903	Novel hum
438	14.4	80.0	3296	12	ADH24264	Novel hum	511	14.4	80.0	3296	12	ADH26800	Novel hum
439	14.4	80.0	3296	12	ADH98618	Novel hum	512	14.4	80.0	3296	12	ADH70463	Novel hum
440	14.4	80.0	3296	12	ADH03449	Novel hum	513	14.4	80.0	3296	12	ADH72209	Novel hum
441	14.4	80.0	3296	12	ADH99170	Novel hum	514	14.4	80.0	3296	12	ADH24078	Novel hum
442	14.4	80.0	3296	12	ADH16755	Novel hum	515	14.4	80.0	3296	12	ADH74539	Novel hum
443	14.4	80.0	3296	12	ADH05214	Novel hum	516	14.4	80.0	3296	12	ADH70464	Novel hum
444	14.4	80.0	3296	12	ADH19481	Novel hum	517	14.4	80.0	3296	12	ADH72197	Novel hum
445	14.4	80.0	3296	12	ADH73563	Novel hum	518	14.4	80.0	3296	12	ADH24080	Novel hum
446	14.4	80.0	3296	12	ADH13318	Novel hum	519	14.4	80.0	3296	12	ADH26810	Novel hum
447	14.4	80.0	3296	12	ADH08375	Novel hum	520	14.4	80.0	3296	12	ADH21451	Novel hum
448	14.4	80.0	3296	12	ADH15845	Novel hum	521	14.4	80.0	3296	12	ADH72195	Novel hum
449	14.4	80.0	3296	12	ADH96943	Novel hum	522	14.4	80.0	3296	12	ADH05323	Novel hum
450	14.4	80.0	3296	12	ADH06128	Novel hum	523	14.4	80.0	3296	12	ADH03470	Novel hum
451	14.4	80.0	3296	12	ADH23712	Novel hum	524	14.4	80.0	3296	12	ADH19174	Novel hum
452	14.4	80.0	3296	12	ADH04001	Novel hum	525	14.4	80.0	3296	12	ADH90226	Novel hum
453	14.4	80.0	3296	12	ADH24902	Novel hum	526	14.4	80.0	3296	12	ADH38855	Novel hum
454	14.4	80.0	3296	12	ADH07719	Novel hum	527	14.4	80.0	3296	12	ADH32815	Novel hum
455	14.4	80.0	3296	12	ADH07751	Novel hum	528	14.4	80.0	3296	12	ADH78433	Novel hum
456	14.4	80.0	3296	12	ADH55246	Novel hum	529	14.4	80.0	3296	12	ADH71542	Novel hum
457	14.4	80.0	3296	12	ADH0910	Novel hum	530	14.4	80.0	3296	12	ADH22303	Novel hum
458	14.4	80.0	3296	12	ADH2014	Novel hum	531	14.4	80.0	3296	12	ADH9682_29	Novel hum
459	14.4	80.0	3296	12	ADH2406	Novel hum	532	14.4	80.0	3296	12	ADH9682_30	Novel hum

Continuation (31 c)

C 533	14.4	80.0	110000	4	AA199683_29	Continuation (30 c	605	13.8	76.7	267	6	AAD43726	Aad43726 Human TR1
534	14.4	80.0	110000	5	AA161373_0	AA161373 Soybean 3	606	13.8	76.7	267	9	ACH666234	ACH666234 Human TR1
535	14.4	80.0	110000	6	ABA90521_11	Continuation (12 c	607	13.8	76.7	267	9	ACH666195	ACH666195 Human TR1
C 536	14.4	80.0	110000	6	ABA90521_11	Continuation (13 c	608	13.8	76.7	267	10	ABX84860	ABX84860 Human TR1
C 537	14.4	80.0	110000	6	ABA90521_13	Continuation (14 c	609	13.8	76.7	273	10	ABX84829	ABX84829 Human TR1
C 538	14.4	80.0	110000	8	ABQ83210_0	Continuation (14 c	610	13.8	76.7	288	10	ABX84829	ABX84829 Human TR1
C 539	14.4	80.0	119596	6	ADH50806	ADH50806 Human G-P	611	13.8	76.7	289	10	ABX84829	ABX84829 Human TR1
540	14.4	80.0	162450	2	AAZ86967	AAZ86967 Retinobla	612	13.8	76.7	301	12	ACH86852	ACH86852 Human TR1
C 541	14.4	77.8	95	10	AAQ28851	AAQ28851 Yeast Chr	613	13.8	76.7	315	8	ABZ52904	ABZ52904 Human TR1
C 542	14.4	77.8	95	10	AAQ28851	AAQ28851 Yeast Chr	614	13.8	76.7	322	8	ACH14949	ACH14949 Human TR1
543	14.4	77.8	256	6	AAAS6967	AAAS6967 Human col	615	13.8	76.7	333	10	ADQ91578	ADQ91578 Human TR1
544	14.4	77.8	256	6	ABT12389	ABT12389 Orestes s	616	13.8	76.7	336	4	AAAS51598	AAAS51598 Human TR1
545	14.4	77.8	256	6	ABT12389	ABT12389 Orestes s	617	13.8	76.7	351	8	ACA466492	ACA466492 Human TR1
C 546	14.4	77.8	552	10	ADP81588	ADP81588 Arabidops	618	13.8	76.7	351	8	ACEF74712	ACEF74712 Human TR1
547	14.4	77.8	552	3	AAFO8080	AAFO8080 Fusarium	619	13.8	76.7	354	4	AAAS55030	AAAS55030 Human TR1
548	14.4	77.8	621	5	ABV86255	ABV86255 Human pro	620	13.8	76.7	354	4	AAAS54523	AAAS54523 Human TR1
549	14.4	77.8	643	8	ABZ52088	ABZ52088 Aspergill	621	13.8	76.7	354	4	ACA20061	ACA20061 Human TR1
550	14.4	77.8	676	3	AAFI1882	AAFI1882 Aspergill	622	13.8	76.7	354	8	ACA20061	ACA20061 Human TR1
551	14.4	77.8	723	11	ABD17176	ABD17176 Pseudom	623	13.8	76.7	391	6	AACT5788	AACT5788 Human TR1
552	14.4	77.8	735	8	ACA43371	ACA43371 Arabidops	624	13.8	76.7	411	4	AAAS60155	AAAS60155 Human TR1
C 553	14.4	77.8	825	11	ABD17306	ABD17306 Pseudom	625	13.8	76.7	456	4	ABZ128221	ABZ128221 Human TR1
C 554	14.4	77.8	847	3	AAAC4129	AAAC4129 Arabidops	626	13.8	76.7	462	2	ABZ128221	ABZ128221 Human TR1
C 555	14.4	77.8	857	6	ABN98841	ABN98841 Arabidops	627	13.8	76.7	462	2	ABZ128221	ABZ128221 Human TR1
C 556	14.4	77.8	1083	11	ABD17277	ABD17277 Pseudom	628	13.8	76.7	462	2	ABZ128221	ABZ128221 Human TR1
C 557	14.4	77.8	1154	3	AAAC51798	AAAC51798 Arabidops	629	13.8	76.7	465	4	AAI134016	AAI134016 Human TR1
C 558	14.4	77.8	1209	6	ABZ33347	ABZ33347 Candida a	630	13.8	76.7	465	4	ABZ24131	ABZ24131 Human TR1
C 559	14.4	77.8	1317	11	ABD17232	ABD17232 Pseudom	631	13.8	76.7	465	4	ABZ24131	ABZ24131 Human TR1
C 560	14.4	77.8	1337	3	AAAC45318	AAAC45318 Arabidops	632	13.8	76.7	465	5	AAI02574	AAI02574 Human TR1
C 561	14.4	77.8	1348	3	AAAC34911	AAAC34911 Arabidops	633	13.8	76.7	465	5	AAI02574	AAI02574 Human TR1
C 562	14.4	77.8	1398	3	AAAC50861	AAAC50861 Arabidops	634	13.8	76.7	465	5	AAI02574	AAI02574 Human TR1
C 563	14.4	77.8	1401	3	AAAC40057	AAAC40057 Arabidops	635	13.8	76.7	477	2	AAI40654	AAI40654 Human TR1
C 564	14.4	77.8	1515	10	ADB31908	ADB31908 DNA encod	636	13.8	76.7	478	6	AAI99070	AAI99070 Human TR1
C 565	14.4	77.8	1515	12	ADDO02122	ADDO02122 Thalecres	637	13.8	76.7	480	12	ADJ44147	ADJ44147 Human TR1
566	14.4	77.8	1621	4	ABLI1465	ABLI1465 Drosophi	638	13.8	76.7	482	3	AAI338133	AAI338133 Human TR1
567	14.4	77.8	1653	8	ACN31082	ACN31082 Prokaryot	639	13.8	76.7	483	4	AAI338133	AAI338133 Human TR1
568	14.4	77.8	1682	4	ABLI1859	ABLI1859 Drosophi	640	13.8	76.7	486	10	ADK53676	ADK53676 Human TR1
C 569	14.4	77.8	1809	4	ABLI28693	ABLI28693 Drosophi	641	13.8	76.7	486	12	ADJ40471	ADJ40471 Human TR1
C 570	14.4	77.8	2058	11	ABD17207	ABD17207 Pseudom	642	13.8	76.7	486	10	ADJ40471	ADJ40471 Human TR1
C 571	14.4	77.8	2210	4	ABLI14469	ABLI14469 Drosophi	643	13.8	76.7	488	3	AAI338133	AAI338133 Human TR1
C 572	14.4	77.8	2324	3	AAAC47073	AAAC47073 Arabidops	644	13.8	76.7	510	12	ABZ53326	ABZ53326 Human TR1
C 573	14.4	77.8	2464	2	AAV03752	AAV03752 Porcine T	645	13.8	76.7	512	12	ACH73121	ACH73121 Human TR1
574	14.4	77.8	2628	3	AACT9600	AACT9600 Vitruene	646	13.8	76.7	520	4	AAI16066	AAI16066 Human TR1
575	14.4	77.8	2628	6	ABQ83477	ABQ83477 Pasteurel	647	13.8	76.7	531	10	ABD03624	ABD03624 Human TR1
576	14.4	77.8	3279	4	ABLI15054	ABLI15054 Drosophi	648	13.8	76.7	546	2	AAI99661	AAI99661 Human TR1
577	14.4	77.8	4384	4	ABLI14468	ABLI14468 Drosophi	649	13.8	76.7	548	2	AAI184270	AAI184270 Human TR1
C 578	14.4	77.8	4448	4	ABLI14412	ABLI14412 Drosophi	650	13.8	76.7	558	6	ABA50401	ABA50401 Human TR1
C 579	14.4	77.8	4715	4	ABLI1858	ABLI1858 Drosophi	651	13.8	76.7	580	6	ABV02684	ABV02684 Human TR1
580	14.4	77.8	4762	4	ABLI18464	ABLI18464 Drosophi	652	13.8	76.7	641	5	ABQ25684	ABQ25684 Human TR1
C 581	14.4	77.8	7025	4	ABLI28692	ABLI28692 Drosophi	653	13.8	76.7	645	6	ABQ25684	ABQ25684 Human TR1
C 582	14.4	77.8	110000	2	AAZ01425_09	Continuation (10 c	654	13.8	76.7	645	6	ABQ25684	ABQ25684 Human TR1
C 583	14.4	77.8	110000	2	ABX16390_0	Continuation (10 c	655	13.8	76.7	650	10	ADK59898	ADK59898 Human TR1
584	14.4	77.8	231222	10	ADLI16393	ADLI16393 Mouse hig	656	13.8	76.7	653	6	ABK74457	ABK74457 Human TR1
585	13.8	76.7	66	2	AAV75954	AAV75954 Stephyloc	657	13.8	76.7	678	3	AAFI13659	AAFI13659 Human TR1
586	13.8	76.7	100	8	ACD78338	ACD78338 E. coli K	658	13.8	76.7	686	6	ABQ124910	ABQ124910 Human TR1
587	13.8	76.7	100	8	ACD78339	ACD78339 E. coli K	659	13.8	76.7	696	6	ABQ124910	ABQ124910 Human TR1
C 588	13.8	76.7	150	8	ACN13639	ACN13639 Prokaryot	660	13.8	76.7	696	6	ABQ124910	ABQ124910 Human TR1
589	13.8	76.7	160	4	AAFI17560	AAFI17560 Human bre	661	13.8	76.7	705	6	ABK74533	ABK74533 Human TR1
590	13.8	76.7	160	4	AAAS46990	AAAS46990 Human bre	662	13.8	76.7	712	4	AAH70542	AAH70542 Human TR1
591	13.8	76.7	160	6	ABK95026	ABK95026 Human bre	663	13.8	76.7	725	9	ADLI14447	ADLI14447 Human TR1
592	13.8	76.7	160	6	ABT08645	ABT08645 Human bre	664	13.8	76.7	728	9	ADLI14447	ADLI14447 Human TR1
593	13.8	76.7	160	6	ABK63591	ABK63591 Human bre	665	13.8	76.7	769	4	AAI37344	AAI37344 Human TR1
594	13.8	76.7	160	10	ABT32803	ABT32803 Human tum	666	13.8	76.7	769	4	ABK60332	ABK60332 Human TR1
595	13.8	76.7	160	10	ADLI29706	ADLI29706 Human bre	667	13.8	76.7	769	12	ADJ31082	ADJ31082 Human TR1
596	13.8	76.7	160	12	ADLI29706	ADLI29706 Human bre	668	13.8	76.7	778	6	ABO89992	ABO89992 Human TR1
597	13.8	76.7	170	8	AAAS50451	AAAS50451 Stephyloc	669	13.8	76.7	783	4	AAAS54375	AAAS54375 Human TR1
C 598	13.8	76.7	170	8	AAAS50451	AAAS50451 Stephyloc	670	13.8	76.7	783	4	AAAS54375	AAAS54375 Human TR1
599	13.8	76.7	223	3	AAAC12283	AAAC12283 Human sec	671	13.8	76.7	801	4	AAAS57018	AAAS57018 Human TR1
600	13.8	76.7	223	3	ABX51442	ABX51442 Bovine ES	672	13.8	76.7	801	10	ADDA2812	ADDA2812 Human TR1
601	13.8	76.7	247	4	AAAS60536	AAAS60536 Human can	673	13.8	76.7	846	8	ACA18772	ACA18772 Human TR1
C 602	13.8	76.7	255	8	ACN14727	ACN14727 Prokaryot	674	13.8	76.7	849	12	ADDO05534	ADDO05534 Human TR1
603	13.8	76.7	267	4	AAI28001	AAI28001 Human TR1	675	13.8	76.7	852	8	ACA51219	ACA51219 Human TR1
604	13.8	76.7	267	4	AAI28040	AAI28040 Human TR1	676	13.8	76.7	852	8	ACA51219	ACA51219 Human TR1
605	13.8	76.7	267	6	AAD43687	AAD43687 Human TR1	677	13.8	76.7	852	8	ACA51219	ACA51219 Human TR1
							678	13.8	76.7	852	10	ADD45795	ADD45795 Human TR1

679	13.8	76.7	869	2	AAX20199	Aax20199	Enterococ	C	752	13.8	76.7	1632	11	ABD13700	Abd13700	Pseudomon
680	13.8	76.7	869	6	ABN98184	Abn98184	E. faecali	C	753	13.8	76.7	1677	10	ADH84959	Adh84959	Enterococ
681	13.8	76.7	869	8	ACA88148	Aca88148	E. faecali	C	754	13.8	76.7	1680	2	AAX20042	Aax20042	Enterococ
682	13.8	76.7	869	8	ABX61754	Abx61754	Enterococ	C	755	13.8	76.7	1680	6	ABN98027	Abn98027	E. faecali
683	13.8	76.7	876	8	ACA42078	Aca42078	Prokaryot	C	756	13.8	76.7	1680	8	ACA87991	Aca87991	E. faecali
684	13.8	76.7	876	8	ABZ15661	Abz15661	Arabiidops	C	757	13.8	76.7	1680	8	ABX61597	Abx61597	Enterococ
685	13.8	76.7	897	6	ADA70428	Ada70428	Rice gene	C	758	13.8	76.7	1680	12	AD061534	Ado61534	Transcrip
686	13.8	76.7	913	5	AAH81625	Aah81625	Human dif	C	759	13.8	76.7	1695	11	ABD13590	Abd13590	Pseudomon
687	13.8	76.7	917	3	AAC45004	Aac45004	Arabiidops	C	760	13.8	76.7	1717	10	ADC03423	Adc03423	Rice flow
688	13.8	76.7	921	11	ABD03698	Abd03698	Pseudomon	C	761	13.8	76.7	1717	12	AAZ41991	Aaz41991	Human end
689	13.8	76.7	924	4	AAAS6173	Aas6173	Salmonell	C	762	13.8	76.7	1719	6	AAAS2047	Aas2047	Parainflu
690	13.8	76.7	924	4	AAI50687	Aai50687	Human pol	C	763	13.8	76.7	1728	10	ADFA4299	Adfa4299	Chinese h
691	13.8	76.7	933	4	ABX80162	Abx80162	Bacillus	C	764	13.8	76.7	1739	8	ABT20494	Abt20494	Aspergill
692	13.8	76.7	933	4	ABX80162	Abx80162	Bacillus	C	765	13.8	76.7	1739	8	ABT18678	Abt18678	Aspergill
693	13.8	76.7	951	5	AAH68382	Aah68382	C. glutam	C	766	13.8	76.7	1749	6	ABZ15564	Abz15564	Arabiidops
694	13.8	76.7	951	10	ACA01298	Aca01298	C. glutam	C	767	13.8	76.7	1758	4	ABL21841	AbL21841	Drosophi
695	13.8	76.7	951	10	ADG88469	Adg88469	Arabiidops	C	768	13.8	76.7	1762	10	ADP82372	Adp82372	Leukemia
696	13.8	76.7	959	6	ABO38182	Abg38182	Oligonuci	C	769	13.8	76.7	1807	4	AAK52630	Aak52630	Human pol
697	13.8	76.7	959	6	ABO38183	Abg38183	Oligonuci	C	770	13.8	76.7	1819	10	ADB94879	Adb94879	Programme
698	13.8	76.7	984	10	ADC93261	Adc93261	E. faecali	C	771	13.8	76.7	1828	4	AAAS26011	Aas26011	Human CDN
699	13.8	76.7	984	4	AAK51646	Aak51646	Human pol	C	772	13.8	76.7	1828	8	ABX73352	Abx73352	Human nov
700	13.8	76.7	999	10	ADH84371	Adh84371	Enterococ	C	773	13.8	76.7	1854	4	AAI58901	Aai58901	Human nov
701	13.8	76.7	1001	3	AAC40948	Aac40948	Arabiidops	C	774	13.8	76.7	1854	5	ADQ99123	Adq99123	DNA encod
702	13.8	76.7	1002	6	AAX20198	Aax20198	Enterococ	C	775	13.8	76.7	1854	6	ADB48883	Adb48883	Novel hum
703	13.8	76.7	1002	8	ACA88147	Aca88147	E. faecali	C	776	13.8	76.7	1874	6	ABX35797	Abx35797	CDNA sequ
704	13.8	76.7	1002	8	ABX61753	Abx61753	Enterococ	C	777	13.8	76.7	1883	12	AD143296	Adi43296	Plant tra
705	13.8	76.7	1006	3	AAQ40943	Aaq40943	Arabiidops	C	778	13.8	76.7	1888	2	AAQ05319	Aaq05319	Viral seq
706	13.8	76.7	1014	12	AD030222	Ado30222	Mouse GPC	C	779	13.8	76.7	1930	5	AAAS92774	Aas92774	DNA encod
707	13.8	76.7	1017	6	ABO70253	Abg70253	Listeria	C	780	13.8	76.7	1949	10	ADC91379	Adc91379	E. faecali
708	13.8	76.7	1041	11	ABD04006	Abd04006	Pseudomon	C	781	13.8	76.7	1965	6	ABO99309	AbO99309	Human cod
709	13.8	76.7	1080	3	AAC53690	Aac53690	Arabiidops	C	782	13.8	76.7	2000	8	ABZ17042	Abz17042	Arabiidops
710	13.8	76.7	1124	2	AAQ82820	Aaq82820	A. aculea	C	783	13.8	76.7	2005	11	ABD13656	Abd13656	Pseudomon
711	13.8	76.7	1125	6	ABZ12421	Abz12421	Arabiidops	C	784	13.8	76.7	2045	2	AAI38260	Aai38260	gene d
712	13.8	76.7	1133	2	AAZ25118	Aaz25118	Corn anth	C	785	13.8	76.7	2055	2	AAI38260	Aai38260	gene d
713	13.8	76.7	1149	8	ACF35366	Acf35366	Mycopacte	C	786	13.8	76.7	2055	3	AAAS1893	Aas1893	T helper
714	13.8	76.7	1149	8	ACA38749	Aca38749	Prokaryot	C	787	13.8	76.7	2055	4	AAAC90981	Aac90981	Mouse ful
715	13.8	76.7	1152	8	ACA40811	Aca40811	Prokaryot	C	788	13.8	76.7	2055	4	AAAF23453	Aaf23453	10 gene n
716	13.8	76.7	1178	6	ABLA9811	AbL9811	Human pol	C	789	13.8	76.7	2055	4	AAI70258	Aai70258	Mouse TH
717	13.8	76.7	1185	5	AAH68383	Aah68383	C. glutami	C	790	13.8	76.7	2055	4	AAAF82604	Aaf82604	Murine TH
718	13.8	76.7	1251	4	AAAS26446	Aas26446	Human CDN	C	791	13.8	76.7	2055	6	AAAS53311	Aas53311	DNA encod
719	13.8	76.7	1251	4	ABX73787	Abx73787	Human nov	C	792	13.8	76.7	2055	9	ABQ77032	Abq77032	Murine T-
720	13.8	76.7	1255	4	AAAS60857	Aas60857	Human can	C	793	13.8	76.7	2055	9	ABQ77032	Abq77032	Murine T-
721	13.8	76.7	1276	3	AAC54737	Aac54737	Arabiidops	C	794	13.8	76.7	2055	11	ADM68617	Adm68617	Different
722	13.8	76.7	1277	3	AAC53414	Aac53414	Arabiidops	C	795	13.8	76.7	2091	9	ADB06513	AdB06513	Alloioioc
723	13.8	76.7	1299	5	ABV28991	Abv28991	Human pro	C	796	13.8	76.7	2091	9	ADB06509	AdB06509	Alloioioc
724	13.8	76.7	1299	5	ABV23150	Abv23150	Human pro	C	797	13.8	76.7	2091	9	ADB06511	AdB06511	Alloioioc
725	13.8	76.7	1299	6	ABZ14067	Abz14067	Arabiidops	C	798	13.8	76.7	2158	4	AAI98108	Aai98108	Human neu
726	13.8	76.7	1299	12	ADN72146	Adn72146	Thale cre	C	799	13.8	76.7	2312	8	ABL24278	AbL24278	Drosophi
727	13.8	76.7	1302	11	ABD13749	Abd13749	Pseudomon	C	800	13.8	76.7	2312	8	ACA27258	Aca27258	Prokaryot
728	13.8	76.7	1325	3	AAC33679	Aac33679	Arabiidops	C	801	13.8	76.7	2425	10	ADP79032	Adp79032	Human pro
729	13.8	76.7	1356	8	AAC51981	Aac51981	Arabiidops	C	802	13.8	76.7	2437	4	AAK94582	Aak94582	Human ful
730	13.8	76.7	1356	8	ACA46498	Aca46498	Prokaryot	C	803	13.8	76.7	2437	12	ADL131475	AdL131475	Pull leng
731	13.8	76.7	1359	4	AAH53220	Aah53220	S. epider	C	804	13.8	76.7	2439	6	ABO99310	AbO99310	Human cod
732	13.8	76.7	1392	6	ABN91628	Abn91628	Staphyloc	C	805	13.8	76.7	2540	4	ABL28220	AbL28220	Drosophi
733	13.8	76.7	1396	3	AAC39628	Aac39628	Arabiidops	C	806	13.8	76.7	2553	9	ACH66191	Ach66191	Human tum
734	13.8	76.7	1399	4	ABL05733	AbL05733	Drosophi	C	807	13.8	76.7	2554	4	AAAF27997	Aaf27997	Human TRI
735	13.8	76.7	1399	4	AAAS7083	Aas7083	CDNA enco	C	808	13.8	76.7	2554	4	AAD43683	Aad43683	Human TRI
736	13.8	76.7	1432	3	AAC42309	Aac42309	Arabiidops	C	809	13.8	76.7	2610	5	AAAF89775	Aaf89775	Nucleotid
737	13.8	76.7	1432	3	AAC42309	Aac42309	Arabiidops	C	810	13.8	76.7	2666	4	ABL26148	AbL26148	Nucleotid
738	13.8	76.7	1435	3	AAC40033	Aac40033	Arabiidops	C	811	13.8	76.7	2733	5	AAAF89774	Aaf89774	Nucleotid
739	13.8	76.7	1440	4	AAAD16713	Aad16713	Cucumis m	C	812	13.8	76.7	2759	4	ABL09909	AbL09909	Drosophi
740	13.8	76.7	1440	4	AAC46559	Aac46559	Zea mays	C	813	13.8	76.7	2779	4	ABL07222	AbL07222	Drosophi
741	13.8	76.7	1470	6	ABZ16057	Abz16057	Arabiidops	C	814	13.8	76.7	2781	2	AAQ13239	Aaq13239	Hsf CDNA
742	13.8	76.7	1502	4	AAD16683	Aad16683	Cucumis m	C	815	13.8	76.7	2781	2	AAQ25712	Aaq25712	Sequence
743	13.8	76.7	1576	3	AAC37497	Aac37497	Arabiidops	C	816	13.8	76.7	2808	10	ADC09472	AdC09472	E. faecali
744	13.8	76.7	1586	4	AAAS15766	Aas15766	Drosophi	C	817	13.8	76.7	2829	4	ABL26387	AbL26387	Drosophi
745	13.8	76.7	1589	12	ADI67069	Adi67069	Novel lac	C	818	13.8	76.7	2895	10	ADA544041	Ada544041	Human cod
746	13.8	76.7	1597	2	ABN20043	Abn20043	Enterococ	C	819	13.8	76.7	2918	4	ABL06526	AbL06526	Drosophi
747	13.8	76.7	1597	6	ABN58028	Abn58028	E. faecali	C	820	13.8	76.7	2936	12	ADK70733	AdK70733	Drosophi
748	13.8	76.7	1597	8	ACA87992	Aca87992	E. faecali	C	821	13.8	76.7	3001	10	ADP25029	AdP25029	Plant gro
749	13.8	76.7	1597	8	ABX61598	Abx61598	Enterococ	C	822	13.8	76.7	3027	12	ADN72586	Adn72586	Thale cre
750	13.8	76.7	1599	8	ABT19272	Abt19272	Aspergill	C	823	13.8	76.7	3096	4	ABL10095	AbL10095	Drosophi
751	13.8	76.7	1599	8	ABT21092	Abt21092	Aspergill	C	824	13.8	76.7	3138	4	ABL08323	AbL08323	Drosophi

825	13.8	76.7	3222	8	ACA25562	Prokaryot	898	13.8	76.7	3501	9	ADA97277	Human PRO
826	13.8	76.7	3280	3	AAA95442	Human CAS	899	13.8	76.7	3501	9	ADB27034	CDNA enco
827	13.8	76.7	3294	4	AAH54973	S. epider	900	13.8	76.7	3501	9	ADB21967	Novel hum
828	13.8	76.7	3331	5	AAF89765	Nucleotid	901	13.8	76.7	3501	9	ADA66658	Human PRO
829	13.8	76.7	3331	10	ADF09659	Human Pre	902	13.8	76.7	3501	9	ADB22519	Human PRO
830	13.8	76.7	3334	4	AAF28030	Human TRL	903	13.8	76.7	3501	9	ADB23392	Human PRO
831	13.8	76.7	3334	6	AAD43716	Human TRL	904	13.8	76.7	3501	9	ADA92014	Novel hum
832	13.8	76.7	3386	4	ACH65224	Human tum	905	13.8	76.7	3501	9	ADBI5077	Human PRO
833	13.8	76.7	3386	4	ACH65177	S. epider	906	13.8	76.7	3501	9	ADBI8329	Human PRO
834	13.8	76.7	3397	10	ADB50622	Human pit	907	13.8	76.7	3501	9	ADB37777	Novel hum
835	13.8	76.7	3429	2	AAQ85413	Human JAK	908	13.8	76.7	3501	10	ADB66249	Novel hum
836	13.8	76.7	3429	3	AAQ66245	Human JAK	909	13.8	76.7	3501	10	ADB89329	Human PRO
837	13.8	76.7	3429	4	ACD03608	Human (Ja	910	13.8	76.7	3501	10	ADB90061	Human PRO
838	13.8	76.7	3429	8	ACC57650	Human pro	911	13.8	76.7	3501	10	ADB39162	Novel hum
839	13.8	76.7	3471	12	ADM79396	Human lym	912	13.8	76.7	3501	10	ADB44785	Novel hum
840	13.8	76.7	3501	4	AA521262	Human CDN	913	13.8	76.7	3501	10	ADB86392	Novel hum
841	13.8	76.7	3501	8	ACA03621	CDNA enco	914	13.8	76.7	3501	10	ADB76997	Novel hum
842	13.8	76.7	3501	8	ABX89159	DNA encod	915	13.8	76.7	3501	10	ADB34154	Human PRO
843	13.8	76.7	3501	8	ACD41813	Human sec	916	13.8	76.7	3501	10	ADB35258	Human PRO
844	13.8	76.7	3501	8	ACA04042	Human CDN	917	13.8	76.7	3501	10	ADB33602	Novel hum
845	13.8	76.7	3501	8	ADA45556	Novel hum	918	13.8	76.7	3501	10	ADB34706	Novel hum
846	13.8	76.7	3501	9	ADA75987	Human PRO	919	13.8	76.7	3501	10	ADB35810	Novel hum
847	13.8	76.7	3501	9	ADA18637	Human PRO	920	13.8	76.7	3501	10	ADB46205	Novel hum
848	13.8	76.7	3501	9	ADA61260	Homo sapi	921	13.8	76.7	3501	10	ADC50078	Novel hum
849	13.8	76.7	3501	9	ADB19045	Novel hum	922	13.8	76.7	3501	10	ADC71625	Novel hum
850	13.8	76.7	3501	9	ADB27586	CDNA enco	923	13.8	76.7	3501	10	ADC59604	Novel hum
851	13.8	76.7	3501	9	ADA86065	Novel hum	924	13.8	76.7	3501	10	ADC52611	Novel hum
852	13.8	76.7	3501	9	ADBI5629	Human PRO	925	13.8	76.7	3501	10	ADC56965	Novel hum
853	13.8	76.7	3501	9	ADA47415	Human PRO	926	13.8	76.7	3501	10	ADC60156	Novel hum
854	13.8	76.7	3501	9	ADA67210	Human PRO	927	13.8	76.7	3501	10	ADC50631	Novel hum
855	13.8	76.7	3501	9	ADB30217	CDNA enco	928	13.8	76.7	3501	10	ADC65158	Novel hum
856	13.8	76.7	3501	9	ADA85513	Novel hum	929	13.8	76.7	3501	10	ADC54256	Novel hum
857	13.8	76.7	3501	9	ADA96725	Human PRO	930	13.8	76.7	3501	10	ADC53217	Novel hum
858	13.8	76.7	3501	9	ADA79029	Human PRO	931	13.8	76.7	3501	10	ADC58740	Novel hum
859	13.8	76.7	3501	9	ADA87168	Novel hum	932	13.8	76.7	3501	10	ADC55618	Novel hum
860	13.8	76.7	3501	9	ADBI6370	Human PRO	933	13.8	76.7	3501	10	ADC58188	Novel hum
861	13.8	76.7	3501	9	ADA91462	Novel hum	934	13.8	76.7	3501	10	ADD002862	Novel hum
862	13.8	76.7	3501	9	ADBI4525	Human PRO	935	13.8	76.7	3501	10	ADC89854	Novel hum
863	13.8	76.7	3501	9	ADBI8486	Novel hum	936	13.8	76.7	3501	10	ADC69273	Novel hum
864	13.8	76.7	3501	9	ADA93701	Human PRO	937	13.8	76.7	3501	10	ADC48162	Human PRO
865	13.8	76.7	3501	9	ADBI9597	Novel hum	938	13.8	76.7	3501	10	ADD09691	Human PRO
866	13.8	76.7	3501	9	ADBI2909	Human PRO	939	13.8	76.7	3501	10	ADD04266	Novel hum
867	13.8	76.7	3501	9	ACD98442	Novel hum	940	13.8	76.7	3501	10	ADC80222	Novel hum
868	13.8	76.7	3501	9	ADA74163	Human PRO	941	13.8	76.7	3501	10	ADD10729	Human PRO
869	13.8	76.7	3501	9	ADB24396	Human PRO	942	13.8	76.7	3501	10	ADC47610	Human PRO
870	13.8	76.7	3501	9	ADA81920	Human PRO	943	13.8	76.7	3501	10	ADC79670	Novel hum
871	13.8	76.7	3501	9	ADA74883	Human PRO	944	13.8	76.7	3501	10	ADD09139	Novel hum
872	13.8	76.7	3501	9	ADA84961	Human PRO	945	13.8	76.7	3501	10	ADD40852	Novel hum
873	13.8	76.7	3501	9	ADA84409	Novel hum	946	13.8	76.7	3501	10	ADD51991	CDNA enco
874	13.8	76.7	3501	9	ADB28665	CDNA enco	947	13.8	76.7	3501	10	ADD54731	CDNA enco
875	13.8	76.7	3501	9	ADA80193	Human PRO	948	13.8	76.7	3501	10	ADD53283	Novel hum
876	13.8	76.7	3501	9	ADA75435	Human PRO	949	13.8	76.7	3501	10	ADD51439	CDNA enco
877	13.8	76.7	3501	9	ADA46660	Human PRO	950	13.8	76.7	3501	10	ADD02238	Human PRO
878	13.8	76.7	3501	9	ADB24956	Human PRO	951	13.8	76.7	3501	10	ADD01672	Human PRO
879	13.8	76.7	3501	9	ADA93132	Human PRO	952	13.8	76.7	3501	10	ADD53854	Novel hum
880	13.8	76.7	3501	9	ADB26482	CDNA enco	953	13.8	76.7	3501	10	ADD92171	Novel hum
881	13.8	76.7	3501	9	ADB30769	CDNA enco	954	13.8	76.7	3501	10	ADD91067	Novel hum
882	13.8	76.7	3501	9	ADA60697	Homo sapi	955	13.8	76.7	3501	10	ADE03681	Novel hum
883	13.8	76.7	3501	9	ADB23844	Human PRO	956	13.8	76.7	3501	10	ADE31978	Novel hum
884	13.8	76.7	3501	9	ADA96173	Human PRO	957	13.8	76.7	3501	10	ADE21910	Novel hum
885	13.8	76.7	3501	9	ADA80745	Human PRO	958	13.8	76.7	3501	10	ADD79134	CDNA enco
886	13.8	76.7	3501	9	ADA95621	Human PRO	959	13.8	76.7	3501	10	ADE41670	Human PRO
887	13.8	76.7	3501	9	ADB25930	CDNA enco	960	13.8	76.7	3501	10	ADE17487	Novel hum
888	13.8	76.7	3501	9	ADB21415	Novel hum	961	13.8	76.7	3501	10	ADD91619	Human PRO
889	13.8	76.7	3501	9	ADA77194	Human PRO	962	13.8	76.7	3501	10	ADD91619	Human PRO
890	13.8	76.7	3501	9	ADBI7934	CDNA enco	963	13.8	76.7	3501	10	ADE33082	Novel hum
891	13.8	76.7	3501	9	ADA86617	Novel hum	964	13.8	76.7	3501	10	ADD79686	CDNA enco
892	13.8	76.7	3501	9	ADA87720	Novel hum	965	13.8	76.7	3501	10	ADD92723	Novel hum
893	13.8	76.7	3501	9	ADA46108	Novel hum	966	13.8	76.7	3501	10	ADE19143	Human PRO
894	13.8	76.7	3501	9	ADB28138	CDNA enco	967	13.8	76.7	3501	10	ADE18591	Human PRO
895	13.8	76.7	3501	9	ADB28690	CDNA enco	968	13.8	76.7	3501	10	ADE42787	Human PRO
896	13.8	76.7	3501	9	ADA76642	Human PRO	969	13.8	76.7	3501	10	ADD95576	Human PRO
897	13.8	76.7	3501	9	ADA88272	Novel hum	970	13.8	76.7	3501	10	ADE22462	CDNA enco

971	13.8	76.7	3501	10	ADD78580	CDNA	encc
972	13.8	76.7	3501	10	AD323530	Novel	hum
973	13.8	76.7	3501	10	AD323530	Novel	hum
974	13.8	76.7	3501	10	AD323530	Novel	hum
975	13.8	76.7	3501	10	AD323530	Novel	hum
976	13.8	76.7	3501	10	AD323530	Novel	hum
977	13.8	76.7	3501	10	AD323530	Novel	hum
978	13.8	76.7	3501	10	AD323530	Novel	hum
979	13.8	76.7	3501	10	AD323530	Novel	hum
980	13.8	76.7	3501	10	AD323530	Novel	hum
981	13.8	76.7	3501	10	AD323530	Novel	hum
982	13.8	76.7	3501	10	AD323530	Novel	hum
983	13.8	76.7	3501	10	AD323530	Novel	hum
984	13.8	76.7	3501	10	AD323530	Novel	hum
985	13.8	76.7	3501	10	AD323530	Novel	hum
986	13.8	76.7	3501	10	AD323530	Novel	hum
987	13.8	76.7	3501	10	AD323530	Novel	hum
988	13.8	76.7	3501	10	AD323530	Novel	hum
989	13.8	76.7	3501	10	AD323530	Novel	hum
990	13.8	76.7	3501	10	AD323530	Novel	hum
991	13.8	76.7	3501	10	AD323530	Novel	hum
992	13.8	76.7	3501	10	AD323530	Novel	hum
993	13.8	76.7	3501	10	AD323530	Novel	hum
994	13.8	76.7	3501	10	AD323530	Novel	hum
995	13.8	76.7	3501	10	AD323530	Novel	hum
996	13.8	76.7	3501	10	AD323530	Novel	hum
997	13.8	76.7	3501	10	AD323530	Novel	hum
998	13.8	76.7	3501	10	AD323530	Novel	hum
999	13.8	76.7	3501	10	AD323530	Novel	hum
1000	13.8	76.7	3501	10	AD323530	Novel	hum

ALIGNMENTS

RESULT 1
Continuation (24 of 24) of ABA90521 from base 2300001 (Genomic sequence of Lactococcus 1
WP Sequence split into 24 fragments LOCUS ABA90521 Accession ABA90521

WP	Fragment Name	Begin	End
WP	ABA90521_00	1	110000
WP	ABA90521_01	100001	210000
WP	ABA90521_02	200001	310000
WP	ABA90521_03	300001	410000
WP	ABA90521_04	400001	510000
WP	ABA90521_05	500001	610000
WP	ABA90521_06	600001	710000
WP	ABA90521_07	700001	810000
WP	ABA90521_08	800001	910000
WP	ABA90521_09	900001	1010000
WP	ABA90521_10	1000001	1110000
WP	ABA90521_11	1100001	1210000
WP	ABA90521_12	1200001	1310000
WP	ABA90521_13	1300001	1410000
WP	ABA90521_14	1400001	1510000
WP	ABA90521_15	1500001	1610000
WP	ABA90521_16	1600001	1710000
WP	ABA90521_17	1700001	1810000
WP	ABA90521_18	1800001	1910000
WP	ABA90521_19	1900001	2010000
WP	ABA90521_20	2000001	2110000
WP	ABA90521_21	2100001	2210000
WP	ABA90521_22	2200001	2310000
WP	ABA90521_23	2300001	2365589

Query Match 91.1%; Score 16.4; DB 6; Length 65589;
Best Local Similarity 94.4%; Pred. No. 4e+02; 1; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACCAACAGATTGC 18
DB 20471 GAGAACCAACAGATTCC 20454

RESULT 2
ABN84784 standard; DNA; 19 BP.
ID ABN84784 standard; DNA; 19 BP.
AC ABN84784;
XX
DT 05-NOV-2002 (first entry)
XX
DE Primer useful for familial dysautonomia allele genotype analysis.
XX
KW Familial dysautonomia; Riley-Day syndrome;
KW hereditary sensory neuropathy III; human; carrier; diagnosis;
KW IkappaB kinase-complex associated protein; primer; SSCP;
KW single-strand conformational polymorphism; ss.
OS Homo sapiens.
XX
PN EP1225232-A2.
XX
PD 24-JUL-2002.
XX
PF 17-FAN-2002; 2002EP-00001232.
XX
PP 17-FAN-2001; 2001US-0262284P.
XX
PR 17-FAN-2001; 2001US-0262284P.
XX
PA (RUBI/ RUBIN B. Y.
PA (ANDE/ ANDERSON S. L.
XX
PI Rubin BY, Anderson SL;
XX
DR WPI; 2002-601228/65.
XX
PT Detecting a polymorphism in a gene encoding the IkappaB kinase-complex-
PT associated protein is used to diagnose and identify carriers of familial
PT dysautonomia.
XX
PS Claim 7; Page 9; 16pp; English.
XX
XX The invention provides a method for detecting a polymorphism linked to a
CC gene associated with familial dysautonomia (FD). This involves detecting
CC a disruptive mutation in a gene encoding the IkappaB kinase-complex-
CC associated protein (IKAP) on chromosome 9q31. Sequence analysis of the
CC IKAP-encoding gene showed, in chromosomes with the major FD haplotype, a
CC T to C transition in position 6 of the donor splice site of intron 20.
CC This mutation (2507+6T to C) results in the generation of an IKAP mRNA in
CC which exon 20 is spliced out along with intron 20. Sequence analysis of
CC the IKAP gene of individuals heterozygous for the FD chromosome with the
CC most common minor haplotype (minor 2) showed a G to C transversion of
CC nucleotide 2390 in exon 19 of the reported IKAP cDNA, resulting in an
CC Arg966Pro amino acid substitution and disruption of a consensus Ser/Thr
CC kinase phosphorylation site. The present sequence is a primer that can be
CC used in a claimed method for detecting a disruptive mutation in the IKAP
CC gene, using single-strand conformational polymorphism (SSCP) analysis.
CC The primer was used in the genotype analysis of FD alleles. Use with the
CC primer given in ABN84784 yielded a 244 bp fragment. In a family with
CC probands homozygous for the major haplotype, all affected individuals
CC were homoallelic for 2507+6C to C and all the parents were heterozygous.
CC In families with probands heterozygous for the major and minor 2 FD
CC haplotypes, 1 parent and the proband were heterozygous for R696P and the
CC other parent and the proband were heterozygous for 2507+6T to C. Analysis
CC of 31 probands homozygous for the major FD haplotype showed that 100% of
CC the probands were homozygous for 2507+6T to C, 100% of the parents were
CC heterozygous for this mutation, and 4 siblings of the probands had FD and
CC were homozygous for the FD haplotype and the 2507+6T to C mutation. Study
CC of a random group of 619 individuals of Ashkenazi Jewish descent revealed
CC the presence of 25 carriers of 2507+6T to C and 2 individuals with R696P.
CC The method is useful for FD diagnosis and for identifying carriers of the
CC condition
XX
SQ Sequence 19 BP; 8 A; 4 C; 4 G; 3 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 19;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
1 GAGAACACACAGATTC 16
Db

RESULT 3

ABQ80567
ID ABQ80567 standard; DNA; 66479 BP.

AC ABQ80567;

DT 08-NOV-2002 (first entry)

DE Mutant human IKEXAP gene #2.

XX Human; IKEXAP; Familial Dysautonomia; FD; Riley-Day syndrome;

KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX FD1; mutation; gene; chromosome 9q31; ds.

OS Homo sapiens.

XX Synthetic.

XX Key

FT mutation Location/Qualifiers
replace(33714,G)
/*tag= a

XX WO200259381-A2.

XX 01-AUG-2002.

XX 07-JAN-2002; 2002WO-US000473.

XX 06-JAN-2001; 2001US-026080P.

XX (GEO) GEN HOSPITAL CORP.

XX Slangenaupt S, Guseella JF;

XX WPI; 2002-674806/72.

XX Claim 1; Page; 109pp; English.

CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKEXAP gene (see ABQ80565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FD1 mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKEXAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKEXAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKEXAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKEXAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKEXAP sequence given in Fig 6

XX Sequence 66479 BP; 18271 A; 12399 C; 14128 G; 21681 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
33979 GAGAACACACAGATTC 33994
Db

RESULT 4

ABQ80566
ID ABQ80566 standard; DNA; 66479 BP.

AC ABQ80566;

DT 08-NOV-2002 (first entry)

DE Mutant human IKEXAP gene #1.

XX Human; IKEXAP; Familial Dysautonomia; FD; Riley-Day syndrome;

KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX FD1; mutation; gene; chromosome 9q31; ds.

OS Homo sapiens.

XX Synthetic.

XX Key

FT mutation Location/Qualifiers
replace(34201,T)
/*tag= a

XX WO200259381-A2.

XX 01-AUG-2002.

XX 07-JAN-2002; 2002WO-US000473.

XX 06-JAN-2001; 2001US-026080P.

XX (GEO) GEN HOSPITAL CORP.

XX Slangenaupt S, Guseella JF;

XX WPI; 2002-674806/72.

XX Claim 1; Page; 109pp; English.

CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKEXAP gene (see ABQ80565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FD1 mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKEXAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKEXAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKEXAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKEXAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKEXAP sequence given in Fig 6

XX Sequence 66479 BP; 18271 A; 12399 C; 14129 G; 21680 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||

Db 33979 GAGAACACACAGATTC 33994

RESULT 5

ABQ80568 standard; DNA; 66479 BP.

XX AC ABQ80568;

XX DT 08-NOV-2002 (first entry)

XX DE Mutant human IKKAP gene #3.

XX XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

XX KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX KM FD1; FD2; mutation; gene; chromosome 9q31; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX FT mutation /tag= a

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

XX FT mutation /tag= b

Sequence 66479 BP; 18271 A; 12400 C; 14128 G; 21680 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||

Db 33979 GAGAACACACAGATTC 33994

RESULT 6

ABQ80565 standard; DNA; 66479 BP.

XX AC ABQ80565;

XX DT 08-NOV-2002 (first entry)

XX DE Human IKKAP wild-type gene.

XX XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

XX KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX KM gene; chromosome 9q31; ds.

XX OS Homo sapiens.

XX OS Homo sapiens.

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

XX FT mutation /tag= a

Sequence 66479 BP; 18271 A; 12398 C; 14129 G; 21681 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||

Db 33979 GAGAACACAGATTC 33994

RESULT 7
ABL37566
ID ABL37566 standard; cDNA, 485 BP.
XX
XX ABL37566;
AC
XX 08-APR-2002 (first entry)
XX
XX Human colon tumour antigen polynucleotide SEQ ID NO:1155.
DE Human colon tumour antigen polynucleotide SEQ ID NO:1155.
XX Human colon tumour antigen; diagnosis; gene; ss.
KW Human colon tumour antigen; diagnosis; gene; ss.
XX Homo sapiens.
OS Homo sapiens.
XX MO200196388-A2.
PN MO200196388-A2.
XX 20-DEC-2001.
BD 20-DEC-2001.
XX 08-JUN-2001; 2001WO-US018557.
PF 08-JUN-2001; 2000US-0210899P.
XX 09-JUN-2000; 2000US-0210899P.
PR 20-FEB-2001; 2001US-0270216P.
XX (CORI-) CORIXA CORP.
PA (CORI-) CORIXA CORP.
XX Jiang Y, Harlocker SL, Secrist H;
PI Jiang Y, Harlocker SL, Secrist H;
XX WPI; 2002-114514/15.
DR WPI; 2002-114514/15.
XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
FT useful for inhibiting development of cancer in patient.
XX Claim 1; SEQ ID NO 1155; 105bp; English.
XX ABL36412 to ABL3645 represent human colon tumour antigen cDNA clones (1)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (1) have cytostatic activity and can be used in vaccine
CC production. (1) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (1) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (1)
CC can be used in the diagnosis of a colon tumour
XX
XX Sequence 485 BP; 121 A; 132 C; 126 G; 104 T; 0 U; 2 Other;
SQ Sequence 485 BP; 121 A; 132 C; 126 G; 104 T; 0 U; 2 Other;
Query Match 85.6%; Score 15.4; DB 6; Length 485;
Best Local Similarity 94.1%; Pred. No. 9e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 AGAACACACAGATTCG 18
Db 48 AGCACACACAGATTCG 64

RESULT 8
ACA32902
ID ACA32902 standard; DNA, 513 BP.
XX
XX ACA32902;
AC
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #14559.
DE Prokaryotic essential gene #14559.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX

OS Enterococcus faecalis.
XX
XX WO200277183-A2.
PN WO200277183-A2.
XX
XX 03-OCT-2002.
PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
PF 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
PR 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
PR 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
PR 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX P-FSDB; AB029032.
DR P-FSDB; AB029032.
XX WPI; 2003-029926/02.
DR WPI; 2003-029926/02.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 20772; 1766bp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 513 BP; 171 A; 81 C; 115 G; 146 T; 0 U; 0 Other;
SQ Sequence 513 BP; 171 A; 81 C; 115 G; 146 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 8; Length 513;
Best Local Similarity 94.1%; Pred. No. 9e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 AGAACACACAGATTCG 18
Db 447 AGAACACACAGATTCG 463

RESULT 9
ADH84638

ID ADH84638 standard; DNA; 534 BP.
XX
AC ADH84638;
XX
DT 22-APR-2004 (first entry)
XX
DE Enterococcus faecalis polynucleotide #2523.
XX
KW Enterococcus faecalis infection; transcription regulatory element;
XX antibacterial; gene; ds.
OS Enterococcus faecalis.
XX US6617156-B1.
XX
XX 09-SEP-2003.
PD
PF 13-AUG-1998; 98US-00134000.
XX
PR 15-AUG-1997; 97US-0055778P.
XX
PA (DOUC/) DOUCETTE-STAMM L A.
XX (BUSH/) BUSH D.
PI Doucette-Stamm LA, Bush D;
XX
DR WPI: 2003-895394/82.
DR P-PSDB; ADH88043.
XX
PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX
PS Disclosure; SEQ ID NO 2523; 193bp; English.
XX
CC The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polynucleotide of the invention.
XX
SQ Sequence 534 BP; 178 A; 82 C; 120 G; 154 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 534;
Best Local Similarity 94.1%; Pred. No. 9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
Db 465 AGAACACACAGATTGCG 481
RESULT 10
ID AAS52481/c
ID AAS52481 standard; DNA; 1470 BP.
XX
AC AAS52481;
XX
DT 13-FEB-2002 (first entry)
XX
DE E. coli DNA for cellular proliferation protein #203.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.
OS Escherichia coli.
XX

PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
PF
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-024578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KJ, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR P-PSDB; AAU34622.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 6118; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae;
CC pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 1470 BP; 279 A; 366 C; 420 G; 405 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
Db 757 AGAACACACAGATTGCG 741
RESULT 11
ID ADE72695
ID ADE72695 standard; DNA; 1822 BP.
XX
AC ADE72695;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human endometrial specific gene, SEQ ID NO 135.
XX
KW cytosolic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
OS Homo sapiens.
XX
XX WO2003060081-A2.
PN

XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002MO-US041612.
XX PR 21-DEC-2001; 2001US-0342756P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Sun Y, Liu C;
XX WPI; 2003-577666/54.
XX DR WPI; 2003-577666/54.
XX PT Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX PS Claim 1; SEQ ID NO 135; 824bp; English.
XX CC The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX SQ Sequence 1822 BP; 312 A; 578 C; 601 G; 331 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 1822;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTCCG 18
Db 952 AGCACACACAGATTCCG 968
RESULT 12
AA88539/c
ID AA88539 standard; cDNA; 1944 BP.
XX AC AA88539;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #24343.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG24352.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 1; SEQ ID NO 24343; 103bp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA64197-AA694564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1944 BP; 419 A; 502 C; 506 G; 517 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ABAACACACAGATTCCG 18
Db 307 AGAACACACAGATTCCG 291
RESULT 13
AA90079/c
ID AA90079 standard; cDNA; 1944 BP.
XX AC AA90079;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #25883.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG25892.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 1; SEQ ID NO 25883; 103bp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 1944 BP; 419 A; 502 C; 506 G; 517 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACAAACAGATTGCG 18
DB 307 AGAACAAACAGATTGCG 291

RESULT 14
AAS93115/C
ID AAS93115 standard; cDNA; 1944 BP.
XX AAS93115;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #28919.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; AAG28928.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 28919; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 1944 BP; 419 A; 502 C; 506 G; 517 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACAAACAGATTGCG 18
DB 307 AGAACAAACAGATTGCG 291

RESULT 15
AAH48474/C
ID AAH48474 standard; DNA; 2186 BP.
XX AAH48474;
XX
XX 21-SEP-2001 (first entry)
XX
XX Escherichia coli lysP gene.
XX
XX L-pipecolic acid production; delta-1-piperidine-6-carboxylic acid;
XX pyrrolidine-5-carboxylate reductase; lysine-6-aminotransferase; lysP, ds.
XX Escherichia coli.
XX
XX WO200148216-A1.
XX
XX 05-JUL-2001.
XX
XX 22-DEC-2000; 2000WO-JP009137.
XX
XX 28-DEC-1999; 99JP-00373389.
XX
XX (SAOC) MERCIAN CORP.
XX
XX Fujii T, Aritoku Y, Mukahara M, Narita T, Agematsu H, Ishiki K;
XX
XX WPI; 2001-418288/44.
XX P-PSDB; AAG64107.
XX
XX Producing L-pipecolic acid for use as a pharmaceutical raw material,
XX comprises reducing delta-1-piperidine-6-carboxylic acid with
XX 5-carboxylate reductase particularly in recombinant bacterium.
XX
XX Disclosure; Page 35-39; 54pp; Japanese.
XX
XX The invention relates to a method for producing L-pipecolic acid. The
XX method comprises reducing delta-1-piperidine-6-carboxylic acid with
XX pyrrolidine-5-carboxylate reductase. A recombinant Escherichia coli or
XX Corynebacterium bacterium that can express a gene encoding lysine-6-
XX aminotransferase is cultured in an L-lysine-containing medium in order to
XX produce L-pipecolic acid. L-pipecolic acid may be used as a
XX pharmaceutical raw material. The process is efficient, and since it is

CC conducted enzymatically, chirality of the final product can be retained.
CC The present sequence is the Escherichia coli lypE gene
XX
SQ Sequence 2186 BP; 444 A; 531 C; 598 G; 613 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 2186;
Best Local Similarity 94.1%; Pred. No. 9.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 1281 AGAACACACAGATTGCG 1265
RESULT 16
ADE72674
ID ADE72674 standard; DNA; 2620 BP.
AC ADE72674;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human endometrial specific gene, SEQ ID NO 114.
XX
XX cytosaratic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 1; SEQ ID NO 114; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
XX Sequence 2620 BP; 523 A; 803 C; 806 G; 488 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 10; Length 2620;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 952 AGAACACACAGATTGCG 968
RESULT 17
ADE72675
ID ADE72675 standard; DNA; 2747 BP.
XX
XX ADE72675;
XX
XX 29-JAN-2004 (first entry)
XX
XX DT

XX Human endometrial specific gene, SEQ ID NO 115.
DE
XX cytosaratic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 1; SEQ ID NO 115; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
XX Sequence 2747 BP; 530 A; 795 C; 836 G; 586 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 10; Length 2747;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 952 AGAACACACAGATTGCG 968
RESULT 18
ADE72693
ID ADE72693 standard; DNA; 2789 BP.
XX
XX ADE72693;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human endometrial specific gene, SEQ ID NO 133.
XX
XX cytosaratic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX DT

XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
PS Claim 1; SEQ ID NO 133; 824bp; English.
XX
CC The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
SQ Sequence 2789 BP; 560 A; 661 C; 839 G; 525 T; 0 U; 4 Other;
XX
Query Match 85.6%; Score 15.4; DB 10; Length 2789;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 AGAACACACAGATTGCG 18
DB 952 AGCACACACAGATTGCG 968
XX
RESULT 19
AAV30824
ID AAV30824 standard; DNA; 3178 BP.
XX
AC AAV30824;
XX
DT 14-SEP-1998 (first entry)
XX
DE Mitochondrial glycerol-3-phosphate dehydrogenase Gnt2 gene.
XX
KM Glycerol-3-phosphate dehydrogenase; G3PDH; GUT2; YII155C; ss.
XX
OS Saccharomyces sp.
XX
FH Key Location/Qualifiers
FT CDS 468..2312
FT /*tag= a
FT /EC_number= "1.1.99.5"
XX
XX WO9821340-A1.
XX
XX 22-MAY-1998.
XX
XX 10-NOV-1997; 97WO-US020293.
XX
XX 13-NOV-1996; 96US-0030602P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (GEMV) GENENCOR INT INC.
XX
XX Balthis BA, Gatenby AA, Haynie SL, Hsu AK, Lareau RD;
XX WPI; 1998-297943/26.
XX P-PSDB; AAW57326.
XX
XX Fermentative production of glycerol using recombinant host - containing
PT genes for glycerol-3-phosphate dehydrogenase and/or glycerol-3-
PT phosphatase.
XX
PS Disclosure; Page 26-27; 57pp; English.
XX
XX This gene, designated GUT2 or YII155C, codes for a mitochondrial glycerol
CC -3-phosphate dehydrogenase (G3PDH, see AAW57327) that catalyses the
CC conversion of dihydroxyacetone phosphate to glycerol-3-phosphate. The
CC invention provides recombinant organisms that express G3PDH and/or
CC glycerol-3-phosphatase (G3P) (see also AAW57324-32) useful for the
CC production of glycerol from a variety of C-sources. A host cell is
CC preferably transformed with a cassette containing a G3PDH gene and/or a
CC G3P gene and then cultured in the presence of a mono-, oligo-,

CC polyaccharide or 1C-substrate. The glycerol obtained is used in
CC cosmetics, liquid soaps, pharmaceuticals, lubricants and antifreezes; its
CC esters are used in the oil and fat industries. The method produces
CC glycerol rapidly and inexpensively without generation of polluting by-
CC products
XX
SQ Sequence 3178 BP; 946 A; 710 C; 646 G; 876 T; 0 U; 0 Other;
XX
Query Match 85.6%; Score 15.4; DB 2; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 AGAACACACAGATTGCG 18
DB 2154 AGAACACACAGATTGCG 2170
XX
RESULT 20
AAV42018
ID AAV42018 standard; DNA; 3178 BP.
XX
AC AAV42018;
XX
DT 12-OCT-1998 (first entry)
XX
DE Glycerol-3-phosphate dehydrogenase gnt2 gene.
XX
KM Glycerol-3-phosphate dehydrogenase; gnt2 gene; 1,3-propanediol; ss.
XX
OS Saccharomyces sp.
XX
FH Key Location/Qualifiers
FT CDS 468..2312
FT /*tag= a
XX
XX WO9821339-A1.
XX
XX 22-MAY-1998.
XX
XX 10-NOV-1997; 97WO-US020292.
XX
XX 13-NOV-1996; 96US-0030601P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (GEMV) GENENCOR INT INC.
XX
XX Gatenby AA, Haynie SL, Nagarajan V, Nair RV, Nakamura CB;
XX Payne MS, Picataggio SK, Dias-Torres W, Hsu AK, Lareau RD;
XX Trimbur DE, Whited GW;
XX WPI; 1998-297942/26.
XX P-PSDB; AAW30682.
XX
XX Fermentative production of 1,3-propanediol - by single organism
PT containing cassette comprising specific genes, and capable of using
PT inexpensive carbon sources.
XX
PS Claim 9; Page 50-52; 95pp; English.
XX
XX This DNA sequence includes the coding region of the gnt2 gene encoding a
CC mitochondrial glycerol-3-phosphate dehydrogenase (see AAW30682) of
CC Saccharomyces sp. This enzyme catalyses the conversion of
CC dihydroxyacetone phosphate to glycerol-3-phosphate. A claimed method for
CC production of 1,3-propanediol (1) comprises culturing a microorganism
CC transformed with a cassette containing at least 1 of the genes (see
CC AAV42012-21) for glycerol-3-phosphate dehydrogenase, glycerol-3-
CC phosphatase, glycerol dehydratase and 1,3-propanediol oxidoreductase (see
CC AAW30686-85). Any of these genes not present on the cassette must be
CC present endogenously. The enzyme sequences may include substitutions,
CC deletions and additions provided activity is not altered. A single
CC recombinant organism can now be used for production of (1) from
CC inexpensive C-sources (contrast use of glycerol or dihydroxyacetone)
CC without causing pollution. (1) is a starting material for polyesters,

CC polyurethanes and cyclic compounds
XX
SQ Sequence 3178 BP; 946 A; 709 C; 647 G; 876 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACAACAAGATTCCG 18
DB 2154 AGAACAACAAGATTCCG 2170
RESULT 21
AAV35739
ID AAV35739 standard; DNA; 3178 BP.
XX
AC AAV35739;
XX
XX 28-SEP-1998 (first entry)
XX
DE Klebsiella pneumoniae glycerol-3-phosphate dehydrogenase GUT2 gene.
XX
KM GUT2 gene; glycerol-3-phosphate dehydrogenase; production;
KM 1,3-propanediol; recombinant; mitochondrial; ss.
XX
OS Klebsiella pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 468..2309
FT /tag= a
FT /product= "glycerol-3-phosphate dehydrogenase"
XX
FN MO9821341-A2.
XX
XX 22-MAY-1998.
XX
PD 13-NOV-1997; 97MO-US020873.
XX
PF 13-NOV-1997; 97MO-US020873.
XX
PR 13-NOV-1996; 96US-0030601P.
XX
PA (GENV) GENENCOR INT INC.
XX
PI Dunn-Coleman NS, Diaz-Torres M, Chase MW, Trimbur D;
XX
DR WPI; 1998-297944/26.
DR P-PSDB; AAW60257.
XX
PT New method for increasing production of 1,3-propane:diol - comprises
PT fermentation of inexpensive carbon sources by microorganism expressing
PT dehydratase, used, e.g. to prolong half-life of enzyme.
XX
PS Disclosure; Page 59-61; 133BP; English.
XX
CC The sequence is that of the GUT2 gene which encodes mitochondrial
CC glycerol-3-phosphate dehydrogenase. It was used as part of a method of
CC fermentative production of 1,3-propanediol (1,3-pd), using an organism
CC comprising at least 1 gene encoding a dehydratase, is improved by
CC inserting into the host a gene encoding protein X and culturing the
CC transformant in presence of a carbon source (e.g. mono-, oligo- or poly-
CC saccharide or 1C substrate) convertible to 1,3-pd. 1,3-pd is a starting
CC material for polyesters, polyurethanes and cyclic compounds. 1,3-pd can
CC now be produced by a single recombinant organism from inexpensive carbon
CC sources such as glucose (rather than costly glycerol or
CC dihydroxyacetone), rapidly and without causing pollution
XX
SQ Sequence 3178 BP; 946 A; 709 C; 647 G; 876 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACAACAAGATTCCG 18

DB 2154 AGAACAACAAGATTCCG 2170
RESULT 22
AA80613
ID AA80613 standard; DNA; 3178 BP.
XX
AC AA80613;
XX
XX 29-SEP-1999 (first entry)
XX
DE GUT2 gene.
XX
XX GUT2, glycerol-3-phosphate dehydrogenase; FAD-dependent enzyme;
KM EC 1.1.99.5; glycerol; recombinant organism; transformation; YII155C;
KM glycerol biosynthetic pathway; expression cassette; 1-3 propanediol;
KM pharmaceutical compound; antifreeze solution; lubricant; polyurethane;
KM cyclic compound; fat and oil industry; polyester fiber; ss.
XX
OS Saccharomyces sp.
XX
FH Key Location/Qualifiers
FT CDS 468..2312
FT /tag= a
FT /product= "glycerol-3-phosphate dehydrogenase"
FT /note= "mitochondrial"
XX
PN MO928480-A1.
XX
XX 10-JUN-1999.
XX
PD 02-DEC-1998; 98MO-US025551.
XX
PF 02-DEC-1997; 97US-00982783.
XX
PR 02-DEC-1997; 97US-00982783.
XX
PA (DUPC) DU PONT DE NEMOURS & CO E I.
XX
PA (GENV) GENENCOR INT INC.
XX
PI Nair RV, Payne MS, Trimbur DE, Valle F;
XX
XX WPI; 1999-385384/32.
XX
DR P-PSDB; AAY26168.
XX
PT Recombinant organisms containing G3PDH and or G3P phosphatase.
XX
PS Disclosure; Page 53-55; 84BP; English.
XX
CC The present sequence is a GUT2 gene alternatively termed YII155C encoding
CC mitochondrial glycerol-3-phosphate dehydrogenase (G3PDH) which catalyses
CC the conversion of dihydroxyacetone phosphate to glycerol-3-phosphate.
CC This is used to produce glycerol from a recombinant organism by
CC transforming a suitable host cell with an expression cassette comprising
CC either one or both of the genes encoding G3PDH and G3P, where the host
CC cell has disruptions in either glycerol kinase or glycerol dehydrogenase
CC endogenous genes to prevent their active expression. The transformed host
CC cell is cultured with a carbon source and glycerol is recovered.
CC Compounds derived from the glycerol biosynthetic pathway like 1,3-
CC propanediol can also be produced. The method provides a rapid,
CC inexpensive and environment-friendly source of glycerol. Glycerol is used
CC in cosmetics, food, pharmaceuticals, lubricants, anti-freeze solutions,
CC fat and oil industry etc. 1,3-propanediol is used for the production of
CC polyester fibers and the manufacture of polyurethanes and cyclic
XX
SQ Sequence 3178 BP; 946 A; 709 C; 647 G; 876 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACAACAAGATTCCG 18

Db 2154 AGAGACAGATTCGC 2170

RESULT 23

ADH72187

ID ADH72187 standard; DNA; 3333 BP.

AC ADH72187;

XX 25-MAR-2004 (first entry)

DE Human gene of the invention NOV51a SEQ ID NO:1083.

XX ds; gene; human; cytosolic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antiparasitic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

XX dyslipidaemia.

CS Homo sapiens.

XX WO2003102155-A2.

PD 11-DEC-2003.

XX 03-JUN-2003; 2003WO-US017430.

PF 03-JUN-2002; 2002US-0385120P.

XX 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 06-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 07-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0387969P.

PR 07-JUN-2002; 2002US-0388681P.

PR 07-JUN-2002; 2002US-0388931P.

PR 07-JUN-2002; 2002US-0389424P.

PR 07-JUN-2002; 2002US-0389717P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389188P.

PR 14-JUN-2002; 2002US-0389170P.

PR 14-JUN-2002; 2002US-0389144P.

PR 14-JUN-2002; 2002US-0389146P.

PR 17-JUN-2002; 2002US-0389729P.

PR 17-JUN-2002; 2002US-0389742P.

PR 18-JUN-2002; 2002US-0389884P.

PR 19-JUN-2002; 2002US-0390006P.

PR 19-JUN-2002; 2002US-0390209P.

PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.

PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402356P.

PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.

PR 12-AUG-2002; 2002US-0402816P.

PR 12-AUG-2002; 2002US-0402821P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.

PR 13-AUG-2002; 2002US-0403531P.

PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403533P.

PR 13-AUG-2002; 2002US-0403563P.

PR 15-AUG-2002; 2002US-0403617P.

PR 15-AUG-2002; 2002US-0403617P.

PR 25-AUG-2002; 2002US-0406182P.

PR 25-AUG-2002; 2002US-0406355P.

PR 27-AUG-2002; 2002US-0406340P.

PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.

PR 23-SEP-2002; 2002US-0412731P.

PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.

PR 30-SEP-2002; 2002US-0414840P.

PR 30-SEP-2002; 2002US-0414954P.

PR 09-OCT-2002; 2002US-0417186P.

PR 09-OCT-2002; 2002US-0417406P.

PR 23-OCT-2002; 2002US-0420639P.

PR 28-OCT-2002; 2002US-0421156P.

PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.

PR 05-NOV-2002; 2002US-00423798.

PR 05-NOV-2002; 2002US-0423798P.

PR 12-NOV-2002; 2002US-0425453P.

XX (CURA-) CURAGEN CORP.

XX Alsbrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;

PI Catterton E, Chappoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;

PI Ettenberg S, Gangoli BA, Gerlach VL, Gorman L, Gunther E, Guo X;

PI Gusev VI, Herrmann UL, Ji W, Kekuda R, Li L, Liu X, MacDougall JR;

PI MacIsachlan T, Malayankar UM, Mezick AJ, Millet I, Mishra VS;

PI Padigaru M, Patirajan M, Pena CE, Peyman JA, Raha D, Raetelli L;

PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shinkets RA;

PI Smithson G, Splek KA, Stone DJ, Vernet CM, Vose EZ, Zhong M;

PI Zhong H;

XX WPI, 2004-081935/08.

DR P-PSDB; ADH72188.

XX New NOVX polypeptides and nucleic acid molecules useful for preventing or

PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or

PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX Example 51; SEQ ID NO 1083; 1880bp; English.

PS The invention relates to a novel isolated polypeptide (NOVX). A

XX polypeptide of the invention has cytosolic, immunomodulator,

CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and

CC antiparasitic activity, and may have a use in gene therapy, and as a

CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the

CC specification. The polypeptide is useful in the manufacture of a

CC medicament for treating a syndrome associated with a human disease. The

CC polypeptide, polynucleotide and antibody are useful in diagnosing,

CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious

CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are

CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence

CC encodes a NOVX polypeptide of the invention.

XX

SQ Sequence 3333 BP; 771 A; 977 C; 936 G; 649 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 3333;

Best Local Similarity 94.1%; Pred. No. 1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAGATTCCG 18
 Db 307 AGCAACAACAGATTCCG 323

RESULT 24
 ADH72193
 ID ADH72193 standard; DNA, 3333 BP.

ADH72193;

25-MAR-2004 (first entry)

Human gene of the invention NOV51d SEQ ID NO:1089.

de; gene; human; cytosstatic; immunomodulator; neuroprotective; nootropic;
 anorectic; antidiabetic; antimicrobial; antilipemic; gene therapy;
 vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
 obesity; diabetes; infectious disease; metabolic syndrome X;
 dyslipidaemia.

Homo sapiens.

MO2003102155-A2.

11-DEC-2003.

03-JUN-2003; 2003WO-US017430.

XX 03-JUN-2002; 2002US-0385120P.
 PR 04-JUN-2002; 2002US-0385784P.
 PR 05-JUN-2002; 2002US-0386041P.
 PR 06-JUN-2002; 2002US-0386047P.
 PR 06-JUN-2002; 2002US-0386376P.
 PR 06-JUN-2002; 2002US-0386453P.
 PR 06-JUN-2002; 2002US-0386864P.
 PR 06-JUN-2002; 2002US-0387016P.
 PR 07-JUN-2002; 2002US-0386796P.
 PR 07-JUN-2002; 2002US-0386816P.
 PR 07-JUN-2002; 2002US-0386931P.
 PR 07-JUN-2002; 2002US-0386942P.
 PR 07-JUN-2002; 2002US-0386971P.
 PR 07-JUN-2002; 2002US-0386976P.
 PR 07-JUN-2002; 2002US-0387262P.
 PR 08-JUN-2002; 2002US-0296960P.
 PR 10-JUN-2002; 2002US-0387400P.
 PR 10-JUN-2002; 2002US-0387535P.
 PR 11-JUN-2002; 2002US-0387625P.
 PR 11-JUN-2002; 2002US-0387633P.
 PR 11-JUN-2002; 2002US-0387668P.
 PR 11-JUN-2002; 2002US-0387696P.
 PR 11-JUN-2002; 2002US-0387702P.
 PR 11-JUN-2002; 2002US-0387836P.
 PR 11-JUN-2002; 2002US-0387859P.
 PR 12-JUN-2002; 2002US-0387933P.
 PR 12-JUN-2002; 2002US-0387934P.
 PR 12-JUN-2002; 2002US-0387960P.
 PR 12-JUN-2002; 2002US-0388022P.
 PR 12-JUN-2002; 2002US-0388096P.
 PR 13-JUN-2002; 2002US-0388123P.
 PR 14-JUN-2002; 2002US-0388118P.
 PR 14-JUN-2002; 2002US-0388129P.
 PR 14-JUN-2002; 2002US-0388144P.
 PR 14-JUN-2002; 2002US-0389146P.
 PR 17-JUN-2002; 2002US-0389729P.
 PR 17-JUN-2002; 2002US-0389742P.
 PR 18-JUN-2002; 2002US-0389844P.
 PR 19-JUN-2002; 2002US-0390066P.
 PR 19-JUN-2002; 2002US-0390099P.
 PR 21-JUN-2002; 2002US-0390763P.
 PR 17-JUL-2002; 2002US-0396706P.
 PR 06-AUG-2002; 2002US-0401628P.
 PR 09-AUG-2002; 2002US-0402156P.
 PR 09-AUG-2002; 2002US-0402256P.
 PR 09-AUG-2002; 2002US-0402389P.
 PR 12-AUG-2002; 2002US-0402786P.
 PR 12-AUG-2002; 2002US-0402815P.
 PR 12-AUG-2002; 2002US-0402821P.
 PR 12-AUG-2002; 2002US-0402832P.
 PR 13-AUG-2002; 2002US-0403448P.
 PR 13-AUG-2002; 2002US-0403459P.
 PR 13-AUG-2002; 2002US-0403531P.
 PR 13-AUG-2002; 2002US-0403532P.
 PR 13-AUG-2002; 2002US-0403563P.
 PR 13-AUG-2002; 2002US-0406317P.
 PR 15-AUG-2002; 2002US-0406317P.
 PR 26-AUG-2002; 2002US-0406182P.
 PR 26-AUG-2002; 2002US-0406355P.
 PR 27-AUG-2002; 2002US-0406240P.
 PR 12-SEP-2002; 2002US-0410084P.
 PR 20-SEP-2002; 2002US-0412528P.
 PR 23-SEP-2002; 2002US-0412731P.
 PR 30-SEP-2002; 2002US-0414801P.
 PR 30-SEP-2002; 2002US-0414839P.
 PR 30-SEP-2002; 2002US-0414840P.
 PR 30-SEP-2002; 2002US-0414954P.
 PR 09-OCT-2002; 2002US-0417186P.
 PR 09-OCT-2002; 2002US-0417406P.
 PR 23-OCT-2002; 2002US-0420639P.
 PR 28-OCT-2002; 2002US-0421156P.
 PR 31-OCT-2002; 2002US-0422690P.
 PR 01-NOV-2002; 2002US-0423130P.
 PR 05-NOV-2002; 2002US-00423798P.
 PR 12-NOV-2002; 2002US-0425453P.
 XX (CURA-) CURAGEN CORP.
 PA Aisbrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
 PI Carterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
 PI Etenberg S, Gangoli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
 PI Gusev V, Hermann JL, Ji W, Kekuda R, Li L, Liu X, MacDougall UB;
 PI MacLachlan T, Malayanar UM, Mezick AV, Mallet T, Mishra VS;
 PI Padigar M, Patirajan M, Pena CE, Payman JA, Rana D, Rascelli L;
 PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shmukets RA;
 PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
 PI Zhong H;
 XX WPI; 2004-061935/08.
 DR P-PSDB; ADH72194.
 DR
 XX New NOVX polypeptides and nucleic acid molecules useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
 PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
 XX
 PS Example 51; SEQ ID NO 1089; 1880bp; English.
 XX
 The invention relates to a novel isolated polypeptide (NOVX). A
 CC polypeptide of the invention has cytosstatic, immunomodulator,
 CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
 CC antilipemic activity, and may have a use in gene therapy, and as a
 CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
 CC any of the 303 fully defined nucleotide sequences given in the
 CC specification. The polypeptide is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease. The
 CC polypeptide, polynucleotide and antibody are useful in diagnosing,
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC encodes a NOVX polypeptide of the invention.
 XX Sequence 3333 BP; 771 A; 978 C; 936 G; 648 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 3333;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
DB 307 AGACACACAGATTGCG 323

RESULT 25
ADE72673

ID ADE72673 standard; DNA; 3368 BP.

AC ADE72673;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 113.

KM cytosstatic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.

OS Homo sapiens.

PN WO2003060081-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US041612.

PR 21-DEC-2001; 2001US-0342756P.

PA (DIAD-) DIADEXUS INC.

PI Sun Y, Liu C;

DR WPI; 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
treating endometrial cancer and non-cancerous disease states in
endometrial.

PS Claim 1; SEQ ID NO 113; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
endometrial specific genes and proteins. The DNA and protein sequences of
the invention are useful for diagnosing, imaging and treating a patient
with endometrial cancer. The present DNA sequence represents a human
endometrial specific gene of the invention.

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

XX Homo sapiens.
OS
PN WO2003060081-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041612.
XX
PR 21-DEC-2001; 2001US-0342756P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C;
XX
DR WPI; 2003-577666/54.
XX
PT Nucleic acid molecules and polypeptides useful for diagnosing and
treating endometrial cancer and non-cancerous disease states in
endometrial.

PS Claim 1; SEQ ID NO 126; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
endometrial specific genes and proteins. The DNA and protein sequences of
the invention are useful for diagnosing, imaging and treating a patient
with endometrial cancer. The present DNA sequence represents a human
endometrial specific gene of the invention.

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

CC CC

PS Claim 1; SEQ ID NO 131; 824bp; English.

XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.

SQ Sequence 3741 BP; 778 A; 1121 C; 1125 G; 717 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3741;

Best Local Similarity 94.1%; Pred. No. 1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18

Db 952 AGCACACACAGATTGCG 968

RESULT 28

ADD69641
ID ADD69641 standard; cDNA; 3858 BP.

AC ADD69641;

DT 15-JAN-2004 (first entry)

DE Human REMAP cDNA - SEQ ID 70.

XX human; receptor and membrane-associated protein; REMAP: cytosolic;
KM antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
KM cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
KM thymoplastic; cell proliferative; cancer; atherosclerosis; neurological;
KM epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS;
KM allergy; developmental; hypothyroidism; Cushing's syndrome; infection;
KM sst gene.

XX Homo sapiens.

XX WO2003048305-A2.

XX 12-JUN-2003.

XX 13-NOV-2002; 2002WO-US036759.

XX 13-NOV-2001; 2001US-0333097P.

XX 15-NOV-2001; 2001US-0335274P.

XX 14-DEC-2001; 2001US-0340542P.

XX 18-DEC-2001; 2001US-0342166P.

XX 11-JAN-2002; 2002US-0347589P.

XX 14-JAN-2002; 2002US-0348687P.

XX (INCY-) INCYTE GENOMICS INC.

XX Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzal Y;

PI Chawla NK, Warren BA, Barroso J, Becha SD, Yue H, Lehr-Mason PM;

PI Thangavelu K, Lee S, Emerling BM, Kaple AE, Khare R, Baughn MR;

PI Gandhi AR, Iran UK, Richardson TW, Marguis JP, Lal PG, Fotsyhe IJ;

PI Lee EA, Swannaker A, Kallick DA, Griffin JA, Elliott VS, Gorvad AB;

PI Harellia ADA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U;

PI Burrill JD, Blake JF, Ho A, Zheng W, Gao J;

XX WPI: 2003-513744/48.
DR P-28DB; ADD69594.

PT New human receptors and membrane-associated proteins (REMAP), useful for
PT diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or
PT infections.

XX Claim 5; SEQ ID NO 70; 298bp; English.

XX The invention relates to a novel isolated polypeptide comprising a human

CC receptor and membrane-associated protein (REMAP). The polypeptide of the
CC invention demonstrates cytostatic, antiarteriosclerotic, anticonvulsant,
CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
CC anti-inflammatory and thymoplastic activities and may be useful for
CC treating and diagnosing various disorders including those which are cell
CC proliferative such as cancer and atherosclerosis, neurological including
CC epilepsy, Huntington's disease and stroke, immune/inflammatory
CC and particularly AIDS and allergies and developmental such as hypothyroidism
CC and Cushing's syndrome, as well as infections. The current sequence is
CC that of the human REMAP cDNA of the invention.

SQ Sequence 3858 BP; 972 A; 961 C; 989 G; 936 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3858;

Best Local Similarity 94.1%; Pred. No. 1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18

Db 368 AGCACACACAGATTGCG 384

RESULT 29

ADE72694
ID ADE72694 standard; DNA; 3886 BP.

AC ADE72694;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 134.

XX cytosolic; vaccine; human; endometrial specific genes;

XX endometrial specific protein; endometrial cancer; ds.

XX Homo sapiens.

XX WO2003060081-A2.

XX 24-JUL-2003.

XX 23-DEC-2002; 2002WO-US041612.

XX 21-DEC-2001; 2001US-0342756P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Liu C;

XX WPI: 2003-577666/54.

XX Nucleic acid molecules and polypeptides useful for diagnosing and

PT treating endometrial cancer and non-cancerous disease states in

PT endometrial.

PS Claim 1; SEQ ID NO 134; 824bp; English.

XX The invention comprises the amino acid and DNA sequences of human

CC endometrial specific genes and proteins. The DNA and protein sequences of

CC the invention are useful for diagnosing, imaging and treating a patient

CC with endometrial cancer. The present DNA sequence represents a human

CC endometrial specific gene of the invention.

SQ Sequence 3886 BP; 846 A; 1090 C; 1123 G; 827 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3886;

Best Local Similarity 94.1%; Pred. No. 1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18

Db 952 AGCACACACAGATTGCG 968

```
RESULT 30
PR ADE72687 standard; DNA; 4067 BP.
XX
XX ADE72687;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific gene, SEQ ID NO 127.
DE
XX cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
PT
XX
XX Claim 1; SEQ ID NO 127; 824bp; English.
PS
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
XX Sequence 4067 BP; 930 A; 1062 C; 1144 G; 931 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 4067;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY
XX 2 AGACACACAGATTGCG 18
XX 952 AGCACAACAGATTGCG 968
DB
XX
XX RESULT 31
XX ADE72692
XX ID ADE72692 standard; DNA; 4331 BP.
XX
XX ADE72692;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific gene, SEQ ID NO 132.
DE
XX cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
PF
```

```
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
PT
XX
XX Claim 1; SEQ ID NO 132; 824bp; English.
PS
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
XX Sequence 4331 BP; 901 A; 1273 C; 1276 G; 881 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 4331;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY
XX 2 AGACACACAGATTGCG 18
XX 952 AGCACAACAGATTGCG 968
DB
XX
XX RESULT 32
XX ADE72671
XX ID ADE72671 standard; DNA; 4404 BP.
XX
XX ADE72671;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific gene, SEQ ID NO 111.
DE
XX
XX cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
PT
XX
XX Claim 1; SEQ ID NO 111; 824bp; English.
PS
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
```


PS Claim 1, SEQ ID NO 112; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.

XX

XX Sequence 4751 BP; 1050 A; 1346 C; 1381 G; 974 T; 0 U; 0 Other;

XX

XX Query Match 85.6%; Score 15.4; DB 10; Length 4751;

XX Best Local Similarity 94.1%; Pred. No. 1e+03;

XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0

XX

Qy 2 AGAACACACAGATTCCG 18
DB 952 AGCACACACAGATTCCG 968

XX

XX RESULT 35
XX ABK99957

XX ID ABK99957 standard; DNA; 4759 BP.

XX AC ABK99957;

XX

XX 21-OCT-2002 (first entry)

XX

XX DNA encoding human secreted protein SCEP-40.

XX

XX Secreted protein; SCEP; human; cell proliferative disorder; cancer;
XX keratosis; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
XX psoriasis; autoimmune disorder; inflammatory disorder; AIDS; arthritis;
XX acquired immunodeficiency syndrome; adult respiratory distress syndrome;
XX Addison's disease; allergy; asthma; osteoporosis; autoimmune thyroiditis;
XX Crohn's disease; dermatitis; diabetes; Graves' disease; haemodialysis;
XX glomerulonephritis; scleroderma; systemic lupus erythematosus; vitellitis;
XX systemic sclerosis; ulcerative colitis; infection; trauma; Pick disease;
XX cardiovascular disorder; congestive heart failure; angina; epilepsy;
XX hypertensive heart disease; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; amyotrophic lateral sclerosis; stroke; dementia;
XX Huntington's disease; multiple sclerosis; neuromuscular disorder;
XX metabolic disorder; endocrine disorder; toxic myopathy; mental disorder;
XX schizophrenic disorder; developmental disorder; anaemia; epilepsy;
XX hypothyroidism; glaucoma; sensorineural hearing loss; cataract;
XX transgenic animal; gene; ds.

XX

XX Homo sapiens.

XX OS

XX PN WO200248337-A2.

XX

XX 20-JUN-2002.

XX

XX 12-DEC-2001; 2001MO-US048517.

XX

XX 13-DEC-2000; 2000US-0255639P.

XX PR 21-DEC-2000; 2000US-0257852P.

XX PR 05-JAN-2001; 2001US-0260105P.

XX PR 18-JAN-2001; 2001US-0262932P.

XX PR 18-JAN-2001; 2001US-0263096P.

XX PR 19-JAN-2001; 2001US-0263090P.

XX PR 02-FEB-2001; 2001US-0265926P.

XX

XX (INCY-) INCYTE GENOMICS INC.

XX

XX Griffin JA, Yao MG, Dugan BM, Yue H, Ding L, Lai PG, Lee EA,
XX Ramkumar J, Thangavelu K, Xu Y, Lee S, Tang YT, Nguyen DB,
XX Marleen BA, Honcheil CD, Gietzen KC, Baughn MR, Gandhi AR, Arvizu C,
XX Malia NK, Lu Y, Elliott VM, Lu DM, Hatalia AJA, Azimzai Y,
XX Khan FA, Tran UK,
XX

XX WPI; 2002-583509/62.
XX PR P-PSDB; ABG69660.

XX

PT	Novel human secreted proteins and polynucleotides for diagnosing,
PT	Preventing or treating cell proliferative, autoimmune/inflammatory,
PT	cardiovascular, neurological and developmental disorders.
XX	
XX	Claim 5; Page 224-225; 234pp; English.
XX	
CC	The invention describes an isolated polypeptide chosen from secreted
CC	proteins (I), SEQ 1-54. (I) and the polynucleotide encoding it (II) are
CC	useful for screening a compound for effectiveness as an agonist or
CC	antagonist of (I) or compound that alters expression of (II). (I), the
CC	identified agonist and antagonist are useful for treating a disease or
CC	condition associated altered expression of functional SECP in a patient.
CC	An antibody specific to (I) is useful for detecting the presence of (I),
CC	purifying (I) from a sample and for diagnosing a condition or disease
CC	associated with expression of SECP in a subject or in a biological
CC	sample. (I) and (II) and modulators of (I) are useful for diagnosis,
CC	treatment and prevention of cell proliferative disorders (e.g. cancer,
CC	keratois, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis and
CC	psoriasis), autoimmune/inflammatory disorders (e.g. acquired
CC	immunodeficiency syndrome (AIDS), adult respiratory distress syndrome,
CC	Addison's disease, allergies, asthma, osteoporosis, autoimmune
CC	thyroiditis, Crohn's disease, dermatitis, diabetes, Graves' disease,
CC	glomerulonephritis, arthritis, scleroderma, systemic lupus erythematosus,
CC	systemic sclerosis, ulcerative colitis, haemodialysis, uveitis, viral,
CC	bacterial, fungal, parasitic, protozoal, helminthic infections and
CC	trauma), cardiovascular disorders (e.g. congestive heart failure, angina,
CC	hypertensive heart disease), neurological disorders (e.g. Alzheimer's and
CC	Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy,
CC	stroke, Huntington's disease, multiple sclerosis, dementia, neuromuscular
CC	disorders, metabolic, endocrine and toxic myopathies, mental disorders,
CC	schizophrenic disorders, and developmental disorders (e.g. anaemia,
CC	epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and
CC	cataract). (II) is useful for creating transgenic animals to model human
CC	disease and to detect and quantify gene expression in biopsied tissues in
CC	which expression of SECP is correlated with disease. This sequence
CC	encodes a human secreted protein (SECP)
CC	
SQ	Sequence 4759 BP; 1176 A; 1226 C; 1257 G; 1100 T; 0 U; 0 Other;
XX	
QY	Query Match 85.6%; Score 15.4; DB 6; Length 4759;
QY	Best Local Similarity 94.1%; Pred. No. 1e+03;
Db	Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
QY	2 AGAACACACAAATTGCC 18
Db	368 AGCACACACAAATTGCC 384
XX	
RESULT 36	
AAAL57275	
ID	AAAL57275 standard; DNA; 4762 BP.
XX	
AC	AAAL57275;
XX	
DT	17-OCT-2003 (first entry)
XX	
DE	LIG-1 'human modifier of p53 pathway' DNA.
XX	
XX	Human modifier; p53 pathway modulating agent; HM; central integrator;
XX	DNA damage; hypoxia; nucleotide deprivation; oncogene activation;
KW	cytostatic; apoptotic; angiogenic; cell proliferation disorder; cancer;
XX	LIG-1; gene; ds; membrane glycoprotein.
XX	
OS	Homo sapiens.
XX	
XX	Key Location/Qualifiers
FT	15..3296
FT	CD5
FT	/*tag= a
FT	/product= "LIG-1 'human modifier of p53 pathway' protein"
XX	
XX	W02003035833-A2.
XX	
XX	01-MAY-2003.
XX	

XX 21-OCT-2002; 2002MO-US033542.
PF 22-OCT-2001; 2001US-0338733P.
XX 15-FEB-2002; 2002US-0357600P.
XX (EXEL-) EXELIXIS INC.
PA Belvin M, Francis-Lang H, Plozman GD, Funke RP, Li D, Friedman L;
PI WPI; 2003-441259/41.
XX P-PSDB; AAC03110.
DR Identifying a candidate p53 pathway modulating agent useful for treating
XX cancer by contacting an assay system comprising a purified HM polypeptide
PT or nucleic acid with a test agent.
XX
XX Example 2; Page 149-152; 205pp; English.
PS The invention relates to a novel method which comprises identifying a
XX candidate p53 pathway-modulating agent. The human modifier (HM)
CC polypeptides and polynucleotides of the invention are identified with the
CC use of Drosophila melanogaster orthologues which are known to modify the
CC p53 pathway. The human p53 protein normally functions as a central
CC integrator of signals including DNA damage, hypoxia, nucleotide
CC deprivation and oncogene activation. The HM molecules of the invention
CC demonstrate cytostatic activity, whilst the method may be useful for
CC identifying a candidate p53 pathway-modulating agent and modulating the
CC p53 pathway in a cell, as well as during the diagnosis and treatment of
CC an apoptotic, angiogenic or cell proliferation disorder, such as cancer.
CC The current sequence is that of the Lig-I membrane glycoprotein HM DNA of
CC the invention
XX
SQ Sequence 4762 BP; 1168 A; 1229 C; 1264 G; 1101 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 8; Length 4762;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Cy 2 AGACACACAGATTCCG 18
Db 382 AGCACACACAGATTCCG 398
RESULT 37
ADH72191
ID ADH72191 standard; DNA; 4762 BP.
XX
AC ADH72191;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NC051c SEQ ID NO:1087.
XX
KW ds; gene; human; cytosolic; immunomodulator; neuroprotective; nocotropic;
KW anorectic; antidiabetic; antipneumonia; antiparasitic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
XX Homo sapiens.
XX OS
XX WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003MO-US017430.
XX
XX 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 06-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0387969P.
PR 07-JUN-2002; 2002US-0388616P.
PR 07-JUN-2002; 2002US-0388931P.
PR 07-JUN-2002; 2002US-0388944P.
PR 07-JUN-2002; 2002US-0388971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388095P.
PR 13-JUN-2002; 2002US-0388124P.
PR 13-JUN-2002; 2002US-0388118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389844P.
PR 19-JUN-2002; 2002US-0390069P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0406317P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-0423798P.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.

(CURA-) CURAGEN CORP.

XX Alsbrook JP, Alvarez E, Anderson DM, Boldog FL, Casman SJ;
PI Cateron E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangoli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Guev VY, Hermann JL, Li W, Kekuda R, Li L, Liu X, Macdonald JR;
PI MacLachlan T, Malyanekar UM, Mezik AJ, Millet I, Mishra VS;
PI Padigara M, Patuturaj M, Pena CE, Peyman J, Raha D, Rastelli L;
PI Rieger DK, Rothberg ME, Sciore P, Shenoy SG, Shinkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CM, Voss EZ, Zhong M;
PI Zhong H;
XX WPI, 2004-081935/08.
DR P-PSDB; ADH72192.
XX
XX
XX New NOXV polypeptides and nucleic acid molecules useful for preventing or
PT treating NOXV-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
XX Example 51: SEQ ID NO 1087, 1860pp; English.
XX
XX The invention relates to a novel isolated polypeptide (NOXV). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOXV polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOXV-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidemias. The nucleic acids are
CC further used as hybridization probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOXV polypeptide of the invention.
XX
XX Sequence 4762 BP, 1168 A; 1229 C; 1264 G; 1101 T; 0 U; 0 Other;
SQ
XX
XX Query March 85.6%; Score 15.4; DB 12; Length 4762;
XX Best Local Similarity 94.18; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACAACAAGATTGCG 18
Db 362 AGCACACACAGATTGCG 398
RESULT 38
ABV93934
ID ABV93934 standard; DNA; 5101 BP.
XX
XX ABV93934;
AC
XX
XX 27-JAN-2003 (first entry)
DT
XX
XX Human NOV30a coding sequence.
DE
XX
XX Human; anti-HIV; cytostatic; antidiabetic; antiparkinsonian; cachexia; AIDS;
KM antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
KM nootropic; immunosuppressive; osteoprotective; antiparkinsonian; cancer;
KM antifertility; cerebroprotective; gene therapy; NOXV; NOV; fertility;
KM metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KM neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KM immune disorder; haematopoietic disorder; cardiovascular disorder;
KM bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KM metabolic syndrome X; wasting disorder; cell differentiation;
KM Single nucleotide polymorphism; SNP; cell proliferation; haematopoiesis;
KM wound healing; angiogenesis; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Variation replace(1871,G)
FT

FT /tag= a
FT /standard name= "Single nucleotide polymorphism"
FT replace(1970,G)
FT /tag= b
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(2001,C)
FT /tag= c
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(2249,T)
FT /tag= d
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(3186,C)
FT /tag= e
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(3318,T)
FT /tag= f
FT /standard name= "Single nucleotide polymorphism"
XX
XX WO200272771-A2.
XX
XX 19-SEP-2002.
PD
XX
XX 06-MAR-2002; 2002MO-US007288.
PF
XX
XX 08-MAR-2001; 2001US-0274101P.
PR
XX 08-MAR-2001; 2001US-0274194P.
PR
XX 08-MAR-2001; 2001US-0274281P.
PR
XX 08-MAR-2001; 2001US-0274322P.
PR
XX 09-MAR-2001; 2001US-0274849P.
PR
XX 12-MAR-2001; 2001US-0275235P.
PR
XX 13-MAR-2001; 2001US-0275579P.
PR
XX 13-MAR-2001; 2001US-0275601P.
PR
XX 14-MAR-2001; 2001US-0276000P.
PR
XX 16-MAR-2001; 2001US-0276776P.
PR
XX 19-MAR-2001; 2001US-0276994P.
PR
XX 20-MAR-2001; 2001US-0277232P.
PR
XX 20-MAR-2001; 2001US-0277321P.
PR
XX 20-MAR-2001; 2001US-0277327P.
PR
XX 20-MAR-2001; 2001US-0277338P.
PR
XX 21-MAR-2001; 2001US-0277791P.
PR
XX 22-MAR-2001; 2001US-0277833P.
PR
XX 23-MAR-2001; 2001US-0278152P.
PR
XX 26-MAR-2001; 2001US-0278894P.
PR
XX 27-MAR-2001; 2001US-0278998P.
PR
XX 27-MAR-2001; 2001US-0279036P.
PR
XX 28-MAR-2001; 2001US-0279344P.
PR
XX 30-MAR-2001; 2001US-0279935P.
PR
XX 30-MAR-2001; 2001US-0280233P.
PR
XX 02-APR-2001; 2001US-0280802P.
PR
XX 02-APR-2001; 2001US-0280832P.
PR
XX 02-APR-2001; 2001US-0280900P.
PR
XX 04-APR-2001; 2001US-0281194P.
PR
XX 13-APR-2001; 2001US-0283675P.
PR
XX 30-APR-2001; 2001US-0287424P.
PR
XX 02-MAY-2001; 2001US-0288066P.
PR
XX 03-MAY-2001; 2001US-0288528P.
PR
XX 15-MAY-2001; 2001US-0281190P.
PR
XX 16-MAY-2001; 2001US-0281099P.
PR
XX 30-MAY-2001; 2001US-0291240P.
PR
XX 31-MAY-2001; 2001US-0294485P.
PR
XX 31-MAY-2001; 2001US-0294889P.
PR
XX 18-JUN-2001; 2001US-0298907P.
PR
XX 19-JUN-2001; 2001US-0299303P.
PR
XX 19-JUN-2001; 2001US-0299310P.
PR
XX 10-JUL-2001; 2001US-0304354P.
PR
XX 31-JUL-2001; 2001US-0309198P.
PR
XX 16-AUG-2001; 2001US-0312903P.
PR
XX 10-SEP-2001; 2001US-0318462P.
PR
XX 12-SEP-2001; 2001US-0318770P.
PR
XX 27-SEP-2001; 2001US-0325430P.
PR

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rng

Page 27

27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-033184P.
PR 14-NOV-2001; 2001US-033272P.
PR 21-NOV-2001; 2001US-0332054P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ,
PI Bollog FL, Li L, Zernusen BD, Tchervnev VT, Gangoli EA, Vernet CAM,
PI Pena CA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK,
PI Voss EZ, Malysankar UM, Anderson DW, Patirajan M, Miller CE,
PI Taupier RJ, Padigarru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF,
PI Zhong M,
XX WPI; 2002-732824/79.
XX P-PSDB; ABB70116.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX
XX Claim 16; Page 208-210; 619pp; English.
XX
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABB70049-ABB70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, hematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
XX Sequence 5101 BP; 1271 A; 1301 C; 1309 G; 1220 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 6; Length 5101;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 2 AGAACAACAATTCGC 18
Db 396 AGCACACAAGATTCGC 412
XX
XX RESULT 39
ADH72189
ID ADH72189 standard; DNA; 5101 BP.
XX
XX AC ADH72189;
XX
XX 25-MAR-2004 (first entry)
XX

DE Human gene of the invention NOV51b SEQ ID NO:1085.
XX
XX ds; gene; human; cytosolic; immunomodulator; neuroprotective; neurotropic;
XX anorectic; antidiabetic; antimicrobial; antipapemic; gene therapy;
XX vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
XX obesity; diabetes; infectious disease; metabolic syndrome X;
XX dyslipidaemia.
XX
XX Homo sapiens.
XX
XX WO2003102155-A2.
XX
XX 11-DEC-2003.
XX
XX 03-JUN-2003; 2003WO-US017430.
XX
XX 03-JUN-2002; 2002US-0385120P.
XX 04-JUN-2002; 2002US-0385784P.
XX 05-JUN-2002; 2002US-0386041P.
XX 05-JUN-2002; 2002US-0386047P.
XX 06-JUN-2002; 2002US-0386376P.
XX 06-JUN-2002; 2002US-0386453P.
XX 06-JUN-2002; 2002US-0386864P.
XX 06-JUN-2002; 2002US-0387016P.
XX 07-JUN-2002; 2002US-0387766P.
XX 07-JUN-2002; 2002US-0387931P.
XX 07-JUN-2002; 2002US-0386931P.
XX 07-JUN-2002; 2002US-0386942P.
XX 07-JUN-2002; 2002US-0386971P.
XX 07-JUN-2002; 2002US-0387262P.
XX 08-JUN-2002; 2002US-0289560P.
XX 10-JUN-2002; 2002US-0387400P.
XX 11-JUN-2002; 2002US-0387555P.
XX 11-JUN-2002; 2002US-0387610P.
XX 11-JUN-2002; 2002US-0387625P.
XX 11-JUN-2002; 2002US-0387634P.
XX 11-JUN-2002; 2002US-0387668P.
XX 11-JUN-2002; 2002US-0387696P.
XX 11-JUN-2002; 2002US-0387702P.
XX 11-JUN-2002; 2002US-0387836P.
XX 11-JUN-2002; 2002US-0387859P.
XX 12-JUN-2002; 2002US-0387933P.
XX 12-JUN-2002; 2002US-0387934P.
XX 12-JUN-2002; 2002US-0387960P.
XX 12-JUN-2002; 2002US-0388022P.
XX 12-JUN-2002; 2002US-0388096P.
XX 12-JUN-2002; 2002US-0389113P.
XX 13-JUN-2002; 2002US-0389118P.
XX 14-JUN-2002; 2002US-0389120P.
XX 14-JUN-2002; 2002US-0389144P.
XX 14-JUN-2002; 2002US-0389146P.
XX 14-JUN-2002; 2002US-0389729P.
XX 17-JUN-2002; 2002US-0389742P.
XX 18-JUN-2002; 2002US-0389844P.
XX 19-JUN-2002; 2002US-0390006P.
XX 19-JUN-2002; 2002US-0390209P.
XX 21-JUN-2002; 2002US-0390763P.
XX 17-JUL-2002; 2002US-0396766P.
XX 06-AUG-2002; 2002US-0401628P.
XX 09-AUG-2002; 2002US-0402156P.
XX 09-AUG-2002; 2002US-0402256P.
XX 09-AUG-2002; 2002US-0402389P.
XX 12-AUG-2002; 2002US-0402786P.
XX 12-AUG-2002; 2002US-0402816P.
XX 12-AUG-2002; 2002US-0402821P.
XX 12-AUG-2002; 2002US-0402832P.
XX 13-AUG-2002; 2002US-0403448P.
XX 13-AUG-2002; 2002US-0403459P.
XX 13-AUG-2002; 2002US-0403531P.
XX 13-AUG-2002; 2002US-0403532P.
XX 13-AUG-2002; 2002US-0403563P.
XX 13-AUG-2002; 2002US-0406317P.
XX 15-AUG-2002; 2002US-0403617P.

PR	26-AUG-2002;	2002US-0406189P.
PR	26-AUG-2002;	2002US-0406355P.
PR	27-AUG-2002;	2002US-0406240P.
PR	12-SEP-2002;	2002US-0410084P.
PR	20-SEP-2002;	2002US-0412528P.
PR	23-SEP-2002;	2002US-0412731P.
PR	30-SEP-2002;	2002US-0414801P.
PR	30-SEP-2002;	2002US-0414839P.
PR	30-SEP-2002;	2002US-0414848P.
PR	30-SEP-2002;	2002US-0414954P.
PR	09-OCT-2002;	2002US-0417186P.
PR	23-OCT-2002;	2002US-0421063P.
PR	28-OCT-2002;	2002US-0422639P.
PR	31-OCT-2002;	2002US-0422690P.
PR	01-NOV-2002;	2002US-0423130P.
PR	05-NOV-2002;	2002US-00423798.
PR	05-NOV-2002;	2002US-0423798P.
PR	12-NOV-2002;	2002US-0425453P.
XX	(CURA-) CURAGEN CORP.	
PI	Alsobrook JP, Alvarez E, Anderson DM, Boldog FL, Casman SU,	
PI	Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K,	
PI	Eitenberg S, Garagollil EA, Gerlach VL, Gorman L, Gunther E, Guo X,	
PI	Gusev VY, Herrmann JH, Ji W, Kekuda R, Li L, Liu X, MacDougall CR,	
PI	Macacchian T, Malyankar UM, Meick AJ, Miller I, Mishra VS,	
PI	Padiagan M, Parturjan M, Pena CEJ, Peyman UB, Raha D, Rastelli L,	
PI	Rieger DK, Rothenberg ME, Sciore P, Shenoy SC, Shimkets RA,	
PI	Smithson G, Spyrek KA, Stone DU, Verner CM, Voss EZ, Zhong M,	
XX	Zhong H;	
XX	WPI; 2004-081935/08.	
DR	P-PsDB; ADH72190.	
XX		
PT	New NOXV polypeptides and nucleic acid molecules useful for preventing or	
PT	treating NOXV-associated disorders, e.g. cancer, diabetes, infection or	
PT	obesity, and in chromosome mapping, tissue typing or pharmacogenomics.	
XX		
PS	Example 51; SEQ ID NO 1085; 1880bp; English.	
XX		
CC	The invention relates to a novel isolated polypeptide (NOXV). A	
CC	polypeptide of the invention has cytostatic, immunomodulator,	
CC	neuroprotective, neurotropic, anorectic, antidiabetic, antimicrobial, and	
CC	antipneumatic activity, and may have a use in gene therapy, and as a	
CC	vaccine. The polypeptides are encoded by NOXV polynucleotides comprising	
CC	any of the 303 fully defined nucleotide sequences given in the	
CC	specification. The polypeptide is useful in the manufacture of a	
CC	medicament for treating a syndrome associated with a human disease. The	
CC	polypeptide, polynucleotide and antibody are useful in diagnosing,	
CC	treating or preventing NOXV-associated disorders, e.g. cancer, cachexia,	
CC	Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious	
CC	diseases, metabolic syndrome x or dyslipidaemias. The nucleic acids are	
CC	further used as hybridisation probes, in chromosome mapping, tissue	
CC	typing, preventive medicine, and pharmacogenomics. The present sequence	
CC	encodes a NOXV polypeptide of the invention.	
XX		
SQ	Sequence 5101 BP; 1271 A; 1301 C; 1309 G; 1220 T; 0 U; 0 Other;	
Query Match	85.6%; Score 15.4; DB 12; Length 5101;	
Best Local Similarity	94.1%; Pred. No. 1e+03;	
Matches 16; Conservative	0; Mismatches 1; Indels 0; Gaps 0	
Oy	2 AGAACCAACAGATTGCG 18	
Dd	396 AGCAACAACAGATTGCG 412	
RESULT 40		
ADE72682		
ID	ADE72682 standard; DNA; 5188 BP.	
XX		
XC	ADE72682;	

[illegible]

XX DBI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX
 PS Claim 1; SEQ ID NO 128; 824pp; English.
 CC The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5204 BP; 1171 A; 1436 C; 1476 G; 1120 T; 0 U; 1 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5204;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGACACACAGATTGCG 18
 DB 952 AGCACAACAGATTGCG 968
 RESULT 42
 ADE72669
 ID ADE72669 standard; DNA; 5263 BP.
 XX ADE72669;
 XX 29-JAN-2004 (first entry)
 XX Human endometrial specific gene, SEQ ID NO 109.
 DE cytostatic; vaccine; human; endometrial specific genes;
 XX endometrial specific protein; endometrial cancer; ds.
 XX Homo sapiens.
 OS
 XX WO2003060081-A2.
 XX 24-JUL-2003.
 XX 23-DEC-2002; 2002WO-US041612.
 XX 21-DEC-2001; 2001US-0342756P.
 PR (DIAD-) DIADEXUS INC.
 PA Sun Y, Liu C;
 PI WPI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX Claim 1; SEQ ID NO 109; 824pp; English.
 CC The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5263 BP; 1211 A; 1412 C; 1485 G; 1155 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5263;
 Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
 DB 952 AGCACAACAGATTGCG 968
 RESULT 43
 ADE72689
 ID ADE72689 standard; DNA; 5273 BP.
 XX ADE72689;
 XX 29-JAN-2004 (first entry)
 XX Human endometrial specific gene, SEQ ID NO 129.
 DE cytostatic; vaccine; human; endometrial specific genes;
 XX endometrial specific protein; endometrial cancer; ds.
 XX Homo sapiens.
 OS
 XX WO2003060081-A2.
 XX 24-JUL-2003.
 XX 23-DEC-2002; 2002WO-US041612.
 XX 21-DEC-2001; 2001US-0342756P.
 PR (DIAD-) DIADEXUS INC.
 PA Sun Y, Liu C;
 PI WPI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX Claim 1; SEQ ID NO 129; 824pp; English.
 PS The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5273 BP; 1216 A; 1419 C; 1482 G; 1156 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5273;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGACACACAGATTGCG 18
 DB 952 AGCACAACAGATTGCG 968
 RESULT 44
 ADN05928
 ID ADN05928 standard; cDNA; 5287 BP.
 XX ADN05928;
 XX 01-JUL-2004 (first entry)
 XX Antipsooriatic cDNA sequence #1198.
 DE ds; gene; antipsooriatic; gene therapy; psoriasis; diagnosis.
 XX Homo sapiens.
 OS
 XX WO2004028479-A2.


```
XX
DT 29-JAN-2004 (first entry)
XX
DE Human endometrial specific gene, SEQ ID NO 107.
XX
KW cytoskeletal; vaccine; human; endometrial specific genes;
KM endometrial specific protein; endometrial cancer; ds.
XX
OS Homo sapiens.
XX
PN WO2003060081-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041612.
XX
PR 21-DEC-2001; 2001US-0342756P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
DR Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
PT Claim 1; SEQ ID NO 107; 824pp; English.
XX
PS The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
SQ Sequence 5353 BP; 1245 A; 1434 C; 1499 G; 1175 T; 0 U; 0 Other;
XX
Query Match 85.6%; Score 15.4; DB 10; Length 5353;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 AGAACAACAGATTGCG 18
DB 952 AGCAACAACAGATTGCG 968
XX
RESULT 48
ADE72681
ID ADE72681 standard; DNA; 5639 BP.
XX
AC ADE72681;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human endometrial specific gene, SEQ ID NO 121.
XX
DE cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
OS Homo sapiens.
XX
PN WO2003060081-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041612.
XX
PR 21-DEC-2001; 2001US-0342756P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C;
```

```
XX
DR WPI; 2003-577666/54.
XX
DE Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
PT Claim 1; SEQ ID NO 121; 824pp; English.
XX
PS The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
SQ Sequence 5639 BP; 1273 A; 1495 C; 1638 G; 1233 T; 0 U; 0 Other;
XX
Query Match 85.6%; Score 15.4; DB 10; Length 5639;
XX Best Local Similarity 94.1%; Pred. No. 1.1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2 AGAACAACAGATTGCG 18
DB 1259 AGCAACAACAGATTGCG 1275
XX
RESULT 49
ADE72685
ID ADE72685 standard; DNA; 5684 BP.
XX
AC ADE72685;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human endometrial specific gene, SEQ ID NO 125.
XX
DE cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
OS Homo sapiens.
XX
PN WO2003060081-A2.
XX
PD 24-JUL-2003.
XX
PF 23-DEC-2002; 2002WO-US041612.
XX
PR 21-DEC-2001; 2001US-0342756P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
DR Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
PT Claim 1; SEQ ID NO 125; 824pp; English.
XX
PS The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
SQ Sequence 5684 BP; 1399 A; 1497 C; 1594 G; 1194 T; 0 U; 0 Other;
XX
Query Match 85.6%; Score 15.4; DB 10; Length 5684;
XX Best Local Similarity 94.1%; Pred. No. 1.1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 2 AGACACACAGATTGGC 18
 |||||
 DB 952 AGACACACAGATTGGC 968

RESULT 50

ADE72670
 ID ADE72670 standard; DNA; 5857 BP.

AC ADE72670;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 110.

KW cytosolic; vaccine; human; endometrial specific genes;
 endometrial specific protein; endometrial cancer; ds.

OS Homo sapiens.

PN WO2003060081-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US041612.

PR 21-DEC-2001; 2001US-0342756P.

FA (DIAD-) DIADEXUS INC.

PI Sun Y, Liu C;

DR WPI; 2803-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
 treating endometrial cancer and non-cancerous disease states in
 endometrial.

PS Claim 1; SEQ ID NO 110; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
 endometrial specific genes and proteins. The DNA and protein sequences of
 the invention are useful for diagnosing, imaging and treating a patient
 with endometrial cancer. The present DNA sequence represents a human
 endometrial specific gene of the invention.

CC Sequence 5857 BP; 1368 A; 1567 C; 1600 G; 1322 T; 0 U; 0 Other;

Query March 85.6%; Score 15.4; DB 10; Length 5857;

Best Local Similarity 94.1%; Pred. No. 1.1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGACACACAGATTGGC 18
 |||||
 DB 952 AGACACACAGATTGGC 968

Search completed: December 3, 2004, 02:24:15
 Job time : 291.211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 23:31:20 ; Search time 563.921 Seconds
(without alignments)
1509.457 Million cell updates/sec

Title: US-10-050-189a-6
Perfect score: 18
Sequence: 1 gagaaacacaaagatcgc 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqe, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : GenEmbl:***
1: gb_ba:***
2: gb_mtg:***
3: gb_in:***
4: gb_cm:***
5: gb_cv:***
6: gb_pat:***
7: gb_ph:***
8: gb_pl:***
9: gb_pt:***
10: gb_ro:***
11: gb_sts:***
12: gb_sy:***
13: gb_un:***
14: gb_vi:***

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	94.4	110000	8	CR382126_06
2	16.4	91.1	10610	1	AE006452
3	16.4	91.1	39228	1	MLCB1788
4	16.4	91.1	100867	8	AC008047
5	16.4	91.1	110000	2	LMFLCHR36_27
6	16.4	91.1	124215	2	AC143118
7	16.4	91.1	142018	8	AY146587
8	16.4	91.1	169261	2	CR352284
9	16.4	91.1	214536	2	CR382377
10	16.4	91.1	239570	2	AC111804
11	16.4	91.1	256123	2	AC111825
12	16.4	91.1	342300	1	MLCPRN8
13	16	88.9	19	6	AY481359
14	16	88.9	64167	2	AC009112
15	16	88.9	66479	6	AX676048
16	16	88.9	78376	9	AL359692
17	16	88.9	88701	8	AC135161
18	16	88.9	110000	8	CR382130_35
19	16	88.9	110000	8	CR382133_06

20	16	88.9	154732	9	AC011700	AC011700 Homo sapi
21	16	88.9	157381	2	AC129818	AC129818 Rattus no
22	16	88.9	168612	2	AC007341	AC007341 Homo sapi
23	16	88.9	180837	9	AC007333	AC007333 Homo sapi
24	16	88.9	214026	2	AC111389	AC111389 Rattus no
25	16	88.9	236172	2	AC094494	AC094494 Rattus no
26	16	88.9	251032	2	AC126106	AC126106 Rattus no
27	16	88.9	292516	2	AC094425	AC094425 Rattus no
28	15.4	88.6	485	6	AX340908	AX340908 Sequence
29	15.4	88.6	534	6	AR396508	AR396508 Sequence
30	15.4	88.6	886	5	AY344560	AY344560 Sinogast
31	15.4	88.6	887	5	AY344557	AY344557 Sinogast
32	15.4	88.6	887	5	AY344559	AY344559 Sinogast
33	15.4	88.6	887	5	AY344561	AY344561 Sinogast
34	15.4	88.6	887	5	AY344562	AY344562 Sinogast
35	15.4	88.6	888	5	AY344558	AY344558 Sinogast
36	15.4	88.6	1524	14	JSRTPMB	JSRTPMB
37	15.4	88.6	1664	8	AK109670	AK109670 Oryza sat
38	15.4	88.6	1950	8	AY692867	AY692867 Saccharom
39	15.4	88.6	1997	8	AB091267	AB091267 Matchanti
40	15.4	88.6	2186	6	BD095617	BD095617 Process f
41	15.4	88.6	2845	6	CO716675	CO716675 Sequence
42	15.4	88.6	3178	6	AR336833	AR336833 Sequence
43	15.4	88.6	3178	6	AR201081	AR201081 Sequence
44	15.4	88.6	3178	6	BD007547	BD007547 Process o
45	15.4	88.6	3178	6	BD009176	BD009176 Method fo
46	15.4	88.6	3178	6	SCGT2A	X71660 S.cerevisia
47	15.4	88.6	3655	1	ECOLYSP	M89774 Escherichia
48	15.4	88.6	3831	2	AB091102	AB091102 Drosophila
49	15.4	88.6	3831	2	AB091102	AB091102 Marchanti
50	15.4	88.6	4717	6	AX463544	AX463544 Sequence
51	15.4	88.6	4759	6	AX463530	AX463530 Sequence
52	15.4	88.6	4762	9	AF381545	AF381545 Homo sapi
53	15.4	88.6	5044	3	AF293351	AF293351 Anopheles
54	15.4	88.6	5287	9	AB050468	AB050468 Homo sapi
55	15.4	88.6	5303	9	BC071561	BC071561 Homo sapi
56	15.4	88.6	6696	4	AF136225	AF136225 Ovis arie
57	15.4	88.6	7043	6	BD193694	BD193694 Enterococ
58	15.4	88.6	7455	14	AF105220	AF105220 Ovine pul
59	15.4	88.6	7462	6	A27950	A27950 JSRV genom
60	15.4	88.6	7462	14	JSRCG	M80216 Jaagsiekte
61	15.4	88.6	7940	1	AF153615	AF153615 Ovis arie
62	15.4	88.6	8924	1	AB014911	AB014911 Streptoco
63	15.4	88.6	10182	1	AB012849	AB012849 Chloxybu
64	15.4	88.6	10472	1	AB015240	AB015240 Shigella
65	15.4	88.6	10474	1	AE005448	AE005448 Escherich
66	15.4	88.6	11791	14	AF357971	AF357971 Ovine pul
67	15.4	88.6	12905	8	AF353329	AF353329 Kallichro
68	15.4	88.6	23223	2	AC005975	AC005975 Drosophila
69	15.4	88.6	28386	3	CEC14A6	Z81470 Caenorhabdi
70	15.4	88.6	30501	3	CER10E12	Z29561 Caenorhabdi
71	15.4	88.6	35412	8	SPC188	AL049662 S.pombe c
72	15.4	88.6	37170	8	CEFC30A10	Z81072 Caenorhabdi
73	15.4	88.6	41803	8	SPC584	AL032824 S.pombe c
74	15.4	88.6	57745	3	CEY6782A	AL132951 Caenorhab
75	15.4	88.6	62355	2	AC084472	AC084472 Caenorhab
76	15.4	88.6	62588	2	AC101236	AC101236 Mus muscu
77	15.4	88.6	67675	2	SC4357	AC021389 Homo sapi
78	15.4	88.6	75317	8	SC4357	AC021389 Homo sapi
79	15.4	88.6	75888	1	ECOHU47	Z38059 S.cerevisia
80	15.4	88.6	75898	14	AY190604	U00007 47 to 48 ce
81	15.4	88.6	75926	10	BX005244	AY190604 Halovirus
82	15.4	88.6	77670	7	AF222060	EX005244 Mouse DNA
83	15.4	88.6	79459	9	AB022217	AF222060 Halovirus
84	15.4	88.6	92602	9	AC010550	AB022217 Arabidops
85	15.4	88.6	99661	2	AC137530	AC010550 Homo sapi
86	15.4	88.6	110000	1	AC137530	AC137530 Takifugu
87	15.4	88.6	110000	2	AC129424_1	Continuation (23 o
88	15.4	88.6	110000	2	AC138525_0	Continuation (23 o
89	15.4	88.6	110000	2	AC138525_1	Continuation (23 o
90	15.4	88.6	110000	2	AC138525_2	Continuation (23 o
91	15.4	88.6	110000	2	CR382128_27	Continuation (12 of
92	15.4	88.6	113566	8	AC002330	Continuation (12 of

93	15.4	85.6	115265	5	AC140941	Gallus ga	C 166	15	83.3	109431	9	AC016696	AC016696 Homo sapi
94	15.4	85.6	123847	8	AC138002	Oryza sat	C 167	15	83.3	110000	8	CR382125_08	Continuation (9 of
95	15.4	85.6	127299	8	AC147434	Medicago	C 168	15	83.3	112328	8	AL390997	Human DNA
96	15.4	85.6	127300	8	AC147429	Medicago	C 169	15	83.3	147706	8	AC083945	AC083945 Oryza sat
97	15.4	85.6	13316	8	AC031985	Homo sapi	C 170	15	83.3	170051	8	AP005772	AP005772 Oryza sat
98	15.4	85.6	139215	8	AC126780	Medicago	C 171	15	83.3	171617	2	AC013616	AC013616 Homo sapi
99	15.4	85.6	146632	2	AC115935	Mus muscu	C 172	15	83.3	173223	2	AP003377	AP003377 Homo sapi
100	15.4	85.6	146773	2	AC013353	Drosophil	C 173	15	83.3	173759	2	CR450686	CR450686 Dario rer
101	15.4	85.6	150339	2	AC141750	Apis mell	C 174	15	83.3	181460	10	AL672181	AL672181 Mouse DNA
102	15.4	85.6	152836	2	CR450718	Danio rer	C 175	15	83.3	181952	2	AC150345	AC150345 Human DNA
103	15.4	85.6	164176	2	AC116409	Mus muscu	C 176	15	83.3	183342	9	AL450345	AL450345 Human DNA
104	15.4	85.6	167627	2	AC079382	Homo sapi	C 177	15	83.3	187265	8	AP004836	AP004836 Oryza sat
105	15.4	85.6	174151	2	AC011182	Homo sapi	C 178	15	83.3	187658	9	AP003366	AP003366 Homo sapi
106	15.4	85.6	174691	9	AC018563	Homo sapi	C 179	15	83.3	192128	2	AC013382	AC013382 Homo sapi
107	15.4	85.6	175448	3	AC099035	Drosophil	C 180	15	83.3	195835	2	AC107840	AC107840 Mus muscu
108	15.4	85.6	181463	9	AC008352	Drosophil	C 181	15	83.3	201210	2	AC024932	AC024932 Homo sapi
109	15.4	85.6	186415	9	AC008675	Homo sapi	C 182	15	83.3	203741	2	AC128206	AC128206 Rattus no
110	15.4	85.6	186418	9	AC018757	Homo sapi	C 183	15	83.3	21855	2	AC106455	AC106455 Rattus no
111	15.4	85.6	189037	2	AC150026	Papio anu	C 184	15	83.3	216402	2	AC099109	AC099109 Rattus no
112	15.4	85.6	189724	2	CR383675	Danio rer	C 185	15	83.3	216680	2	AC021211	AC021211 Homo sapi
113	15.4	85.6	189827	5	BR890610	Zebrafish	C 186	15	83.3	222680	2	AC110091	AC110091 Mus muscu
114	15.4	85.6	192063	9	AC099665	Homo sapi	C 187	15	83.3	229315	2	AC105651	AC105651 Rattus no
115	15.4	85.6	196491	9	AC092832	Homo sapi	C 188	15	83.3	233157	2	AC105468	AC105468 Rattus no
116	15.4	85.6	199615	8	ATCHR1V7	Arabidops	C 189	15	83.3	235045	2	AC119588	AC119588 Rattus no
117	15.4	85.6	203107	8	AC139255	Homo sapi	C 190	15	83.3	237080	2	AC106251	AC106251 Rattus no
118	15.4	85.6	205764	2	AC079478	Mus muscu	C 191	15	83.3	238597	2	AC131857	AC131857 Rattus no
119	15.4	85.6	213514	2	AC079500	Mus muscu	C 192	15	83.3	238722	2	AC098808	AC098808 Rattus no
120	15.4	85.6	213514	2	AC079500	Mus muscu	C 193	15	83.3	239760	2	AC126206	AC126206 Rattus no
121	15.4	85.6	214429	10	AL671925	Mouse DNA	C 194	15	83.3	240479	2	AC096377	AC096377 Rattus no
122	15.4	85.6	222629	2	AC125802	Rattus no	C 195	15	83.3	274786	2	AC109661	AC109661 Rattus no
123	15.4	85.6	225158	5	AL953901	Zebrafish	C 196	15	83.3	300029	8	AE017083	AE017083 Oryza sat
124	15.4	85.6	225654	5	AL953842	Zebrafish	C 197	15	83.3	348402	2	AC120619	AC120619 Rattus no
125	15.4	85.6	230188	2	AC091604	Mus muscu	C 198	14.8	82.2	294	11	G71609	G71609 A62987734FW
126	15.4	85.6	236612	2	AC111126	Mus muscu	C 199	14.8	82.2	304	11	G71923	G71923 A6496634FW
127	15.4	85.6	245071	2	AC134078	Rattus no	C 200	14.8	82.2	320	11	G70842	G70842 A62987734FB
128	15.4	85.6	245558	2	AC079514	Mus muscu	C 201	14.8	82.2	400	11	SV133261	SV133261 Bos tauru
129	15.4	85.6	248381	2	AC009937	Homo sapi	C 202	14.8	82.2	467	6	AX585805	AX585805 Sequence
130	15.4	85.6	254250	1	AP005350	Vibrio vu	C 203	14.8	82.2	467	6	AX585805	AX585805 Sequence
131	15.4	85.6	265053	1	AP008057	Rattus no	C 204	14.8	82.2	468	11	SV143259	SV143259 P203027 Z
132	15.4	85.6	268687	1	AP002560	Escherich	C 205	14.8	82.2	471	11	SV143266	SV143266 P203027 Z
133	15.4	85.6	276991	2	AC129258	Rattus no	C 206	14.8	82.2	481	11	SV143267	SV143267 P203027 Z
134	15.4	85.6	283700	3	AE003459	Drosophill	C 207	14.8	82.2	483	11	SV143268	SV143268 P203027 Z
135	15.4	85.6	295689	1	AE016985	Shigella	C 208	14.8	82.2	488	11	SV143268	SV143268 P203027 Z
136	15.4	85.6	300029	1	AE016952	Enterococ	C 209	14.8	82.2	488	11	SV143266	SV143266 P203027 Z
137	15.4	85.6	300150	1	AP004598	Oceanobac	C 210	14.8	82.2	489	11	SV143266	SV143266 P203027 Z
138	15.4	85.6	300964	8	AE016890	Eremothec	C 211	14.8	82.2	491	6	BD178253	BD178253 Mold poly
139	15.4	85.6	301130	1	AE016763	Escherich	C 212	14.8	82.2	491	6	BD179332	BD179332 Polynucleo
140	15.4	85.6	301332	1	AE017012	Bacillus	C 213	14.8	82.2	491	11	SV143270	SV143270 P203027 Z
141	15.4	85.6	306803	1	AE017161	Prochloro	C 214	14.8	82.2	493	11	SV143264	SV143264 P203027 Z
142	15.4	85.6	312163	2	AC126132	Rattus no	C 215	14.8	82.2	493	11	SV143265	SV143265 P203027 Z
143	15.4	85.6	313573	3	CEY57611C	Caenorhabd	C 216	14.8	82.2	493	11	SV143267	SV143267 P203027 Z
144	15.4	85.6	319485	2	AC078979	Homo sapi	C 217	14.8	82.2	495	11	SV143268	SV143268 P203027 Z
145	15.4	85.6	330724	2	CEY67H2	Caenorhab	C 218	14.8	82.2	495	11	SV143271	SV143271 P203027 Z
146	15	83.3	485	6	BD262112	Composite	C 219	14.8	82.2	495	11	SV143266	SV143266 P203027 Z
147	15	83.3	485	6	AR360897	Sequence	C 220	14.8	82.2	569	11	SV000188	SV000188 S209P6007
148	15	83.3	737	8	BT000168	Arabidops	C 221	14.8	82.2	578	8	AY299269	AY299269 Arabidops
149	15	83.3	744	11	PM2G9G	Penicilliu	C 222	14.8	82.2	581	11	DMVATP14	DMVATP14 D.melanogas
150	15	83.3	869	8	AY136338	Arabidops	C 223	14.8	82.2	585	11	G99063	G99063 S208P6545RB
151	15	83.3	891	8	AY087026	Arabidops	C 224	14.8	82.2	602	6	CO587817	CO587817 Sequence
152	15	83.3	1455	6	AX488833	Sequence	C 225	14.8	82.2	634	6	AX721295	AX721295 Sequence
153	15	83.3	1689	9	PAURA3	Y09921 P. anom	C 226	14.8	82.2	700	14	HAU22898	HAU22898 Human adeno
154	15	83.3	36798	9	HSE78G1	Human DNA	C 227	14.8	82.2	722	8	AY299268	AY299268 Arabidops
155	15	83.3	42781	8	MISOLSUR	Scenedesmus	C 228	14.8	82.2	759	11	BV036917	BV036917 Arabidops
156	15	83.3	42939	8	AF204057	Scenedesmus	C 229	14.8	82.2	813	11	BV062894	BV062894 Arabidops
157	15	83.3	44499	8	AB016876	Arabidops	C 230	14.8	82.2	840	8	BT012634	BT012634 Arabidops
158	15	83.3	66240	2	AC090603	Homo sapi	C 231	14.8	82.2	952	8	AB126868	AB126868 Aspergill
159	15	83.3	81609	8	AC027035	Arabidops	C 232	14.8	82.2	993	8	GN132218	GN132218 Gnetum gn
160	15	83.3	84162	6	T22J18	Sequence	C 233	14.8	82.2	1143	10	AF155474	AF155474 Sigmodon
161	15	83.3	95769	6	AX652138	Genomic s	C 234	14.8	82.2	1143	10	AF296188	AF296188 Sigmodon
162	15	83.3	95769	6	AC013430	Arabidops	C 235	14.8	82.2	1143	10	AF425215	AF425215 Sigmodon
163	15	83.3	97242	8	AC051630	Arabidops	C 236	14.8	82.2	1143	10	AF425216	AF425216 Sigmodon
164	15	83.3	107816	8	AC016447	Arabidops	C 237	14.8	82.2	1143	10	AF425217	AF425217 Sigmodon
165	15	83.3	109027	9	AC096717	Homo sapi	C 238	14.8	82.2	1202	8	AY485263	AY485263 Zea mays

C 239	14.8	82.2	1253	8	AY452809	AY452809 Fusarium	C 312	14.8	82.2	17716	3	U20862	U20862 Caenorhabdi
C 240	14.8	82.2	1292	8	BT236854	AJ236854 Bos tauru	C 313	14.8	82.2	20704	2	LMFLCHR5_7	Continuation (8 of
C 241	14.8	82.2	1336	8	AF212615	AF212615 Fusarium	C 314	14.8	82.2	20702	1	AE008869	AE008869 Salmonell
C 242	14.8	82.2	1336	8	AF212616	AF212616 Fusarium	C 315	14.8	82.2	23424	2	AC017219	AC017219 Drosophi
C 243	14.8	82.2	1336	8	AF212617	AF212617 Fusarium	C 316	14.8	82.2	24186	3	CE005H5	CE005H5 Caenorhabdi
C 244	14.8	82.2	1336	8	AF212618	AF212618 Fusarium	C 317	14.8	82.2	24342	2	CE049A5	CE049A5 Caenorhabdi
C 245	14.8	82.2	1336	6	AX078926	AX078926 Sequence	C 318	14.8	82.2	29216	3	AC136556	AC136556 Rattus no
C 246	14.8	82.2	1377	4	AF036953	AF036953 Felis cat	C 319	14.8	82.2	30352	3	U28412	U28412 Caenorhabdi
C 247	14.8	82.2	1377	4	AF036953	AF036953 Felis cat	C 320	14.8	82.2	35708	2	AX084516	AX084516 Sequence
C 248	14.8	82.2	1517	8	AK109902	AK109902 Oryza sat	C 321	14.8	82.2	35724	6	AX084519	AX084519 Sequence
C 249	14.8	82.2	1552	8	AT082203	AT082203 Arabidopsi	C 322	14.8	82.2	36114	6	AY183926	AY183926 Drosophi
C 250	14.8	82.2	1564	8	AY096554	AY096554 Arabidops	C 323	14.8	82.2	39125	3	AY183926	AY183926 Drosophi
C 251	14.8	82.2	1692	8	AY224467	AY224467 Oryza sat	C 324	14.8	82.2	39235	3	CE046B3	CE046B3 Caenorhabdi
C 252	14.8	82.2	1821	4	BOVSL37B	L11063 Bos tauru	C 325	14.8	82.2	39694	2	AC144439	AC144439 Streptomy
C 253	14.8	82.2	1830	9	BOVSL37B	BOVSL37B Homo sapi	C 326	14.8	82.2	41097	1	AF016585	AF016585 Caenorhabdi
C 254	14.8	82.2	1863	6	AY056415	AY056415 Arabidops	C 327	14.8	82.2	41396	9	HSU73633	HSU73633 Rattus no
C 255	14.8	82.2	2000	6	AX656526	AX656526 Sequence	C 328	14.8	82.2	42845	9	U88315	U88315 Caenorhabdi
C 256	14.8	82.2	2139	3	AY058544	AY058544 Drosophi	C 329	14.8	82.2	43721	3	U88315	U88315 Caenorhabdi
C 257	14.8	82.2	2157	6	CQ588615	CQ588615 Sequence	C 330	14.8	82.2	44380	1	AE001137	AE001137 Borrelia
C 258	14.8	82.2	2401	6	AK065899	AK065899 Oryza sat	C 331	14.8	82.2	46264	2	AC137490	AC137490 Homo sapi
C 259	14.8	82.2	2529	6	CQ804028	CQ804028 Sequence	C 332	14.8	82.2	47000	8	AP005896	AP005896 Oryza sat
C 260	14.8	82.2	2560	6	AK111850	AK111850 Oryza sat	C 333	14.8	82.2	47452	3	CE045N09	CE045N09 Caenorhab
C 261	14.8	82.2	2602	6	CQ587816	CQ587816 Sequence	C 334	14.8	82.2	48505	2	AC137489	AC137489 Arabidops
C 262	14.8	82.2	2645	14	AY144570	AY144570 Rice blic	C 335	14.8	82.2	50641	8	AB008270	AB008270 Arabidops
C 263	14.8	82.2	2645	14	RBL409148	AJ409148 Rice blic	C 336	14.8	82.2	53797	9	AP001415	AP001415 Homo sapi
C 264	14.8	82.2	2727	6	AR320492	AR320492 Sequence	C 337	14.8	82.2	53905	5	AL359292	AL359292 Human DNA
C 265	14.8	82.2	2821	8	AF233592	AF233592 Arabidops	C 338	14.8	82.2	53905	5	CR388122	CR388122 Zebrafish
C 266	14.8	82.2	2859	8	AK099981	AK099981 Oryza sat	C 339	14.8	82.2	53905	5	AC100232	AC100232 Drosophi
C 267	14.8	82.2	3030	6	AC014405	AC014405 Drosophi	C 340	14.8	82.2	60729	2	AC100232	AC100232 Drosophi
C 268	14.8	82.2	3081	6	AK777288	AK777288 Sequence	C 341	14.8	82.2	66241	2	AC019757	AC019757 Homo sapi
C 269	14.8	82.2	3098	6	CQ589908	CQ589908 Sequence	C 342	14.8	82.2	72828	2	AC026849	AC026849 Arabidops
C 270	14.8	82.2	3098	6	CQ589908	CQ589908 Sequence	C 343	14.8	82.2	73617	3	AC018226	AC018226 Drosophi
C 271	14.8	82.2	3141	3	DMDWMT3MR	X647736 D.meianogae	C 344	14.8	82.2	74901	3	AC005474	AC005474 Arabidops
C 272	14.8	82.2	3169	8	AF188994	AF188994 Arabidops	C 345	14.8	82.2	76339	8	AB0059246	AB0059246 Homo sapi
C 273	14.8	82.2	3181	3	AF289998	AF289998 Drosophi	C 346	14.8	82.2	77939	8	AB0059246	AB0059246 Homo sapi
C 274	14.8	82.2	3213	3	AY122223	AY122223 Drosophi	C 347	14.8	82.2	79303	9	ATC31P16	ATC31P16 Lotus cor
C 275	14.8	82.2	3231	10	AK120280	AK120280 Oryza sat	C 348	14.8	82.2	79516	8	ATC31P16	ATC31P16 Lotus cor
C 276	14.8	82.2	3231	10	MMU487561	AJ487561 Mus muscu	C 349	14.8	82.2	80088	9	AP004909	AP004909 Arabidops
C 277	14.8	82.2	3242	1	MTDNAHEWA	X33691 M.thermoaut	C 350	14.8	82.2	82389	8	AC113430	AC113430 Homo sapi
C 278	14.8	82.2	3444	10	MMU487960	AJ487960 Mus muscu	C 351	14.8	82.2	82406	9	AC026315	AC026315 Homo sapi
C 279	14.8	82.2	3511	6	CQ585996	CQ585996 Sequence	C 352	14.8	82.2	83495	2	AC020432	AC020432 Arabidops
C 280	14.8	82.2	3511	6	AX777374	AX777374 Sequence	C 353	14.8	82.2	83495	2	AP003739	AP003739 Rattus no
C 281	14.8	82.2	3516	6	AY051763	AY051763 Drosophi	C 354	14.8	82.2	86358	2	ATF23E12	ATF23E12 Rattus no
C 282	14.8	82.2	3521	6	CQ588952	CQ588952 Sequence	C 355	14.8	82.2	86710	8	AC098131	AC098131 Oryza sat
C 283	14.8	82.2	3693	6	AK102211	AK102211 Oryza sat	C 356	14.8	82.2	92034	2	AP003770	AP003770 Arabidops
C 284	14.8	82.2	3693	6	CQ600860	CQ600860 Sequence	C 357	14.8	82.2	92187	8	AC024685	AC024685 Homo sapi
C 285	14.8	82.2	3862	3	DROWN75A	M97450 Drosophi	C 358	14.8	82.2	93629	2	AC142017	AC142017 Rattus no
C 286	14.8	82.2	3864	3	BT010268	BT010268 Drosophi	C 359	14.8	82.2	93629	2	AC142017	AC142017 Rattus no
C 287	14.8	82.2	4624	6	CQ603234	CQ603234 Sequence	C 360	14.8	82.2	97837	8	AP004794	AP004794 Arabidops
C 288	14.8	82.2	4624	6	AF444780	AF444780 Anopheles	C 361	14.8	82.2	99220	8	AP004794	AP004794 Arabidops
C 289	14.8	82.2	5631	3	CQ595920	CQ595920 Sequence	C 362	14.8	82.2	99241	9	AP0079030	AP0079030 Homo sapi
C 290	14.8	82.2	5824	6	CQ585995	CQ585995 Sequence	C 363	14.8	82.2	100000	9	AP000017	AP000017 Homo sapi
C 291	14.8	82.2	6367	5	AC125231	AC125231 Homo sapi	C 364	14.8	82.2	100000	9	AP0000159	AP0000159 Homo sapi
C 292	14.8	82.2	6368	9	AX344615	AX344615 Sequence	C 365	14.8	82.2	100725	2	AC149889	AC149889 Xenopus t
C 293	14.8	82.2	6636	8	AY170127	AY170127 Arabidops	C 366	14.8	82.2	102543	5	HS22109	HS22109 Rattus no
C 294	14.8	82.2	6711	8	CQ589907	CQ589907 Sequence	C 367	14.8	82.2	103039	8	AC091292	AC091292 Oryza sat
C 295	14.8	82.2	7146	6	BACTEUBAC	CQ577344 Sequence	C 368	14.8	82.2	104470	8	AP004137	AP004137 Oryza sat
C 296	14.8	82.2	7442	6	CQ577344	CQ577344 Sequence	C 369	14.8	82.2	104470	8	AP004137	AP004137 Oryza sat
C 297	14.8	82.2	8043	6	CQ583253	CQ583253 Sequence	C 370	14.8	82.2	106159	4	ATC42041	ATC42041 Rattus no
C 298	14.8	82.2	8231	6	CQ586520	CQ586520 Sequence	C 371	14.8	82.2	106159	4	ATC42041	ATC42041 Rattus no
C 299	14.8	82.2	8731	6	AE001797	AE001797 Thermocog	C 372	14.8	82.2	107972	2	AP003839	AP003839 Oryza sat
C 300	14.8	82.2	9404	1	AC020107	AC020107 Drosophi	C 373	14.8	82.2	107972	2	AE000516	AE000516 Sequence
C 301	14.8	82.2	9712	2	AE012811	AE012811 Chlorobiu	C 374	14.8	82.2	110000	1	AE016822	AE016822 Oryza sat
C 302	14.8	82.2	10029	2	AE012811	AE012811 Chlorobiu	C 375	14.8	82.2	110000	1	AE016822	AE016822 Oryza sat
C 303	14.8	82.2	10419	2	AE012811	AE012811 Chlorobiu	C 376	14.8	82.2	110000	2	AC098564	AC098564 Homo sapi
C 304	14.8	82.2	10443	1	AE013914	AE013914 Homo sapi	C 377	14.8	82.2	110000	2	AC098564	AC098564 Homo sapi
C 305	14.8	82.2	10615	1	AE012316	AE012316 Xanthomon	C 378	14.8	82.2	110000	2	AP006496	AP006496 Oryza sat
C 306	14.8	82.2	12326	5	AY779214	AY779214 Rivulif m	C 379	14.8	82.2	110000	2	AP006496	AP006496 Oryza sat
C 307	14.8	82.2	14050	1	AE000731	AE000731 Aquilif a	C 380	14.8	82.2	110000	2	EX890596	EX890596 Homo sapi
C 308	14.8	82.2	14350	1	AE013856	AE013856 Yersinia	C 381	14.8	82.2	110000	2	EX890596	EX890596 Homo sapi
C 309	14.8	82.2	15437	1	AF002133	AF002133 Mycobacte	C 382	14.8	82.2	110000	2	CEY105E8	CEY105E8 Rattus no
C 310	14.8	82.2	15684	1	AB002150	AB002150 Bacillus	C 383	14.8	82.2	110000	2	LMFLCHR15_6	Continuation (7 of
C 311	14.8	82.2	15683	1	AE001748	AE001748 Thermocog	C 384	14.8	82.2	110000	2	LMFLCHR32_00	Continuation (7 of

385	14.8	82.2	110000	2	LMFLCHR36.01	Continuation (2 of	458	14.8	82.2	157381	2	AC129818	AC129818 Rattus no
386	14.8	82.2	110000	2	LMFLCHR36.06	Continuation (7 of	459	14.8	82.2	157442	2	AC129829	AC129829 Homo sapi
387	14.8	82.2	110000	2	LMFLCHR36.31	Continuation (32 of	460	14.8	82.2	157460	2	AC129838	AC129838 Homo sapi
388	14.8	82.2	110000	2	CR380954.06	Continuation (7 of	461	14.8	82.2	158241	2	AC021804	AC021804 Mus muscu
389	14.8	82.2	110000	8	CR382132.22	Continuation (22 o	462	14.8	82.2	158296	2	BX530098	BX530098 Homo sapi
390	14.8	82.2	110000	8	CR382132.35	Continuation (36 o	463	14.8	82.2	158971	2	AP004882	AP004882 Homo sapi
391	14.8	82.2	110000	8	AC145127_15	Continuation (16 o	464	14.8	82.2	159148	2	AP005801	AP005801 Oryza sat
392	14.8	82.2	112155	8	CR378661	AC092145 Homo sapi	465	14.8	82.2	161144	9	AC074142	AC074142 Homo sapi
393	14.8	82.2	113176	2	AC092145	AC092145 Homo sapi	466	14.8	82.2	161273	9	AL160171	AL160171 Homo sapi
394	14.8	82.2	113382	2	AC149886	AC149886 Xenopus t	467	14.8	82.2	161547	9	AP001929	AP001929 Homo sapi
395	14.8	82.2	114288	2	AP000590	AP000590 Homo sapi	468	14.8	82.2	161902	2	CR450715	CR450715 Homo sapi
396	14.8	82.2	117555	8	AC104285	AC104285 Oryza sat	469	14.8	82.2	161905	10	AL929034	AL929034 Mouse DNA
397	14.8	82.2	118192	8	AP004120	AP004120 Oryza sat	470	14.8	82.2	162017	9	AC093522	AC093522 Homo sapi
398	14.8	82.2	119388	8	CNS0808C8	EX569688 Oryza sat	471	14.8	82.2	162046	9	AC068776	AC068776 Homo sapi
399	14.8	82.2	120391	8	CNS0808C8	AL318176 Oryza sat	472	14.8	82.2	162449	2	AP004377	AP004377 Oryza sat
400	14.8	82.2	120787	8	AC000292	AC000292 Arabidops	473	14.8	82.2	164449	2	AC092320	AC092320 Homo sapi
401	14.8	82.2	121521	2	EX572074	EX572074 Dario rer	474	14.8	82.2	165378	2	CR381620	CR381620 Homo sapi
402	14.8	82.2	123384	2	EX649279	EX649279 Zebrafish	475	14.8	82.2	165652	2	AC068010	AC068010 Homo sapi
403	14.8	82.2	124918	10	AC140843	AC140843 Mus muscu	476	14.8	82.2	165931	2	AC124032	AC124032 Rattus no
404	14.8	82.2	126062	8	AP006135	AP006135 Oryza sat	477	14.8	82.2	166046	2	AC021808	AC021808 Homo sapi
405	14.8	82.2	126062	2	AC078977	AC078977 Oryza sat	478	14.8	82.2	167292	9	AL157713	AL157713 Human DNA
406	14.8	82.2	127313	2	AC099527	AC099527 Felis cat	479	14.8	82.2	167300	10	AL928678	AL928678 Mouse DNA
407	14.8	82.2	127341	2	AC099527	AL513467 Neurospor	480	14.8	82.2	167830	9	AC016549	AC016549 Homo sapi
408	14.8	82.2	130049	8	NC1785	AL513467 Neurospor	481	14.8	82.2	167871	5	EX649473	EX649473 Zebrafish
409	14.8	82.2	131182	10	AC132255	AL844513 Mouse DNA	482	14.8	82.2	168583	3	AC104510	AC104510 Drosophill
410	14.8	82.2	131184	2	AL844513	AL844513 Mouse DNA	483	14.8	82.2	168583	9	AL354920	AL354920 Human DNA
411	14.8	82.2	132060	2	AC138555	AC138555 Carolinia	484	14.8	82.2	169123	2	CR376728	CR376728 Dario rer
412	14.8	82.2	132729	8	AP005849	AP005849 Oryza sat	485	14.8	82.2	169123	2	AC069164	AC069164 Homo sapi
413	14.8	82.2	133309	2	AL355003	AL355003 Human DNA	486	14.8	82.2	169903	2	AC023712	AC023712 Drosophill
414	14.8	82.2	134978	2	AC079004	AC079004 Homo sapi	487	14.8	82.2	170129	2	AC023712	AC023712 Drosophill
415	14.8	82.2	135226	2	AC148970	AC148970 Strongylio	488	14.8	82.2	170761	2	AC129115	AC129115 Rattus no
416	14.8	82.2	135294	8	AC105320	AC105320 Oryza sat	489	14.8	82.2	170896	10	AC140199	AC140199 Mus muscu
417	14.8	82.2	137659	8	AL139232	AL139232 Human DNA	490	14.8	82.2	170956	2	AC127007	AC127007 Rattus no
418	14.8	82.2	139630	10	AC139635	AC139635 Mus muscu	491	14.8	82.2	171032	2	AC091761	AC091761 Felis cat
419	14.8	82.2	140282	8	AC108504	AC108504 Oryza sat	492	14.8	82.2	171695	2	AC012337	AC012337 Homo sapi
420	14.8	82.2	140948	2	AC091795	AC091795 Felis cat	493	14.8	82.2	172153	2	AC145287	AC145287 Xenopus t
421	14.8	82.2	141384	2	AL358194	AL358194 Homo sapi	494	14.8	82.2	172437	2	AC022078	AC022078 Homo sapi
422	14.8	82.2	141926	2	AC140687	AC140687 Rattus no	495	14.8	82.2	173165	3	AC099005	AC099005 Drosophill
423	14.8	82.2	143336	2	HS1106N8	AL035457 Human DNA	496	14.8	82.2	173192	8	EX927386	EX927386 Dario rer
424	14.8	82.2	147764	2	AC026119	AC026119 Homo sapi	497	14.8	82.2	173522	2	OSJN00060	OSJN00060 Oryza sat
425	14.8	82.2	148298	9	AC022081	AC022081 Homo sapi	498	14.8	82.2	174028	8	AC087363	AC087363 Homo sapi
426	14.8	82.2	149298	8	OSJN00114	AL066994 Oryza sat	499	14.8	82.2	174314	3	AC011066	AC011066 Drosophill
427	14.8	82.2	149498	2	AC019348	AC019348 Homo sapi	500	14.8	82.2	174735	3	AC104449	AC104449 Homo sapi
428	14.8	82.2	149501	2	AC127682	AC127682 Mus muscu	501	14.8	82.2	176329	2	AC006402	AC006402 Drosophill
429	14.8	82.2	145415	5	AC147826	AC147826 Xenopus t	502	14.8	82.2	176802	2	AC012383	AC012383 Homo sapi
430	14.8	82.2	146497	2	AC147204	AC147204 Xenopus t	503	14.8	82.2	176802	2	AC113177	AC113177 Mus muscu
431	14.8	82.2	146505	8	AP004299	AP004299 Oryza sat	504	14.8	82.2	176992	2	AC010959	AC010959 Homo sapi
432	14.8	82.2	147123	8	AP005691	AP005691 Oryza sat	505	14.8	82.2	177672	2	AC113343	AC113343 Rattus no
433	14.8	82.2	147983	8	AP018981	AC118981 Genomic s	506	14.8	82.2	177962	10	AC099640	AC099640 Mus muscu
434	14.8	82.2	149148	9	AC108460	AC108460 Homo sapi	507	14.8	82.2	177962	2	AC016705	AC016705 Homo sapi
435	14.8	82.2	149259	9	AC099654	AC099654 Homo sapi	508	14.8	82.2	178420	2	EX571836	EX571836 Dario rer
436	14.8	82.2	149369	10	AL607123	AL607123 Mouse DNA	509	14.8	82.2	178420	2	AC015524	AC015524 Homo sapi
437	14.8	82.2	149496	2	AC053499	AC053499 Homo sapi	510	14.8	82.2	179015	2	AC135839	AC135839 Bos tauru
438	14.8	82.2	150192	9	AC008686	AC008686 Homo sapi	511	14.8	82.2	179859	2	AC007563	AC007563 Homo sapi
439	14.8	82.2	150474	8	AP003635	AP003635 Oryza sat	512	14.8	82.2	179859	8	AP003827	AP003827 Oryza sat
440	14.8	82.2	151094	2	AC091797	AC091797 Felis cat	513	14.8	82.2	179898	2	AC009338	AC009338 Mus muscu
441	14.8	82.2	151163	8	AC079128	AC079128 Oryza sat	514	14.8	82.2	180395	2	AC008280	AC008280 Homo sapi
442	14.8	82.2	151514	5	AL953907	AL953907 Zebrafish	515	14.8	82.2	180395	2	AC021436	AC021436 Homo sapi
443	14.8	82.2	151759	2	AC149967	AC149967 Strongylio	516	14.8	82.2	180765	9	AC007521	AC007521 Drosophill
444	14.8	82.2	152355	9	AC005668	AC005668 Homo sapi	517	14.8	82.2	180919	2	AC007251	AC007251 Homo sapi
445	14.8	82.2	152772	9	AC009326	AC009326 Homo sapi	518	14.8	82.2	180964	9	AC011065	AC011065 Drosophill
446	14.8	82.2	153148	2	AC017030	AC017030 Homo sapi	519	14.8	82.2	181103	3	EX548162	EX548162 Dario rer
447	14.8	82.2	153452	2	AC044895	AC044895 Homo sapi	520	14.8	82.2	181134	2	AC147355	AC147355 Xenopus t
448	14.8	82.2	153455	2	AC145260	AC145260 Felis cat	521	14.8	82.2	181256	2	AC147169	AC147169 Homo sapi
449	14.8	82.2	153508	2	AC141790	AC141790 Apis mell	522	14.8	82.2	181604	2	AC090442	AC090442 Homo sapi
450	14.8	82.2	153535	2	AC090429	AC090429 Homo sapi	523	14.8	82.2	182408	2	AL135872	AL135872 Homo sapi
451	14.8	82.2	153792	2	AC099049	AC099049 Oryza sat	524	14.8	82.2	182450	3	AC103548	AC103548 Drosophill
452	14.8	82.2	154161	8	AC131967	AC131967 Oryza sat	525	14.8	82.2	182901	2	AC093198	AC093198 Homo sapi
453	14.8	82.2	154656	8	AP004341	AP004341 Oryza sat	527	14.8	82.2	182941	2	AC118882	AC118882 Rattus no
454	14.8	82.2	154859	9	AP0069045	AP0069045 Homo sapi	528	14.8	82.2	184444	5	EX0050471	EX0050471 Zebrafish
455	14.8	82.2	155006	2	AC141681	AC141681 Apis mell	529	14.8	82.2	184499	2	AC104910	AC104910 Mus muscu
456	14.8	82.2	155587	8	AP005156	AP005156 Oryza sat	530	14.8	82.2	184705	2	AC104910	AC104910 Mus muscu
457	14.8	82.2	156244	8	CNS0771X	AL627332 DNA cent							

C 531	14.8	82.2	185330	2	AC147410	Homo sapi	604	14.8	82.2	227370	2	AC114867	AC114867	Rattus no
C 532	14.8	82.2	185790	3	AC091221	Drosophila	C 605	14.8	82.2	227854	2	AC122584	AC122584	Rattus no
C 533	14.8	82.2	186104	9	AP005718	Homo sapi	C 606	14.8	82.2	228331	2	AC096430	AC096430	Rattus no
C 534	14.8	82.2	186341	9	AC023351	Homo sapi	C 607	14.8	82.2	228978	2	AC096427	AC096427	Rattus no
C 535	14.8	82.2	186517	2	AC116386	Mus muscu	C 608	14.8	82.2	229076	2	AC107142	AC107142	Rattus no
C 536	14.8	82.2	186541	2	AC017057	Homo sapi	C 609	14.8	82.2	229247	10	AC098888	AC098888	Mus muscu
C 537	14.8	82.2	187555	3	AC010037	Drosophila	C 610	14.8	82.2	229726	2	AC095735	AC095735	Rattus no
C 538	14.8	82.2	188489	3	AC012098	Drosophila	C 611	14.8	82.2	229872	10	AL591514	AL591514	Mouse DNA
C 539	14.8	82.2	188716	5	AL807739	Zebrafish	C 612	14.8	82.2	230526	2	AC114457	AC114457	Rattus no
C 540	14.8	82.2	190833	5	BX005041	Zebrafish	C 613	14.8	82.2	231254	5	BX784023	BX784023	Zebrafish
C 541	14.8	82.2	191046	10	AL607146	Mouse DNA	C 614	14.8	82.2	231499	2	AC111629	AC111629	Rattus no
C 542	14.8	82.2	191598	5	BX537308	Danio rer	C 615	14.8	82.2	231679	2	AC123319	AC123319	Rattus no
C 543	14.8	82.2	191785	5	BX005113	Zebrafish	C 616	14.8	82.2	232936	2	AC105642	AC105642	Rattus no
C 544	14.8	82.2	191851	2	AC149099	Rattus no	C 617	14.8	82.2	233000	2	AC146694	AC146694	Rattus no
C 545	14.8	82.2	193371	2	AC098928	Rattus no	C 618	14.8	82.2	233422	10	AC123675	AC123675	Rattus no
C 546	14.8	82.2	194301	2	AP000896	Homo sapi	C 619	14.8	82.2	234431	2	AC113307	AC113307	Rattus no
C 547	14.8	82.2	195651	5	BX294114	Zebrafish	C 620	14.8	82.2	236281	2	AC113341	AC113341	Rattus no
C 548	14.8	82.2	196115	2	AC148460	Xenopus t	C 621	14.8	82.2	236783	2	AC106064	AC106064	Rattus no
C 549	14.8	82.2	196217	2	CR388195	Danio rer	C 622	14.8	82.2	239701	2	AC094922	AC094922	Rattus no
C 550	14.8	82.2	196428	2	AC148723	Salmo sal	C 623	14.8	82.2	239893	2	AC107521	AC107521	Rattus no
C 551	14.8	82.2	197424	2	AC148616	Salmo sal	C 624	14.8	82.2	239940	2	AC123752	AC123752	Mus muscu
C 552	14.8	82.2	197859	8	ATCRRIV83	Alu61587 Arabidops	C 625	14.8	82.2	240483	2	AC127628	AC127628	Rattus no
C 553	14.8	82.2	197890	2	AC149092	Par trogl	C 626	14.8	82.2	240729	2	AC111961	AC111961	Rattus no
C 554	14.8	82.2	200105	2	AP002175	Homo sapi	C 627	14.8	82.2	241178	2	AC130508	AC130508	Rattus no
C 555	14.8	82.2	201156	2	AC021757	Homo sapi	C 628	14.8	82.2	243361	2	AC094132	AC094132	Rattus no
C 556	14.8	82.2	201320	2	AC146868	Xenopus t	C 629	14.8	82.2	243846	2	AC128864	AC128864	Rattus no
C 557	14.8	82.2	201322	9	AC091194	Homo sapi	C 630	14.8	82.2	246237	3	CEY33864	CEY33864	Rattus no
C 558	14.8	82.2	201328	2	AC141866	Rattus no	C 631	14.8	82.2	246852	2	AC115675	AC115675	Rattus no
C 559	14.8	82.2	202039	10	AL844182	Mouse DNA	C 632	14.8	82.2	249682	2	AC098402	AC098402	Rattus no
C 560	14.8	82.2	202328	2	AC148617	Salmo sal	C 633	14.8	82.2	250865	5	BX248410	BX248410	Zebrafish
C 561	14.8	82.2	202330	9	AC055872	Homo sapi	C 634	14.8	82.2	251301	2	AC094403	AC094403	Rattus no
C 562	14.8	82.2	202360	2	AC074594	Homo sapi	C 635	14.8	82.2	252395	2	AC130744	AC130744	Rattus no
C 563	14.8	82.2	202891	2	AC103407	Mus muscu	C 636	14.8	82.2	252985	2	AC110727	AC110727	Mus muscu
C 564	14.8	82.2	202876	2	AC010413	Homo sapi	C 637	14.8	82.2	253440	2	AC0960347	AC0960347	Vbrito vu
C 565	14.8	82.2	203876	2	AC128278	Rattus no	C 638	14.8	82.2	254650	1	AC0905347	AC0905347	Rattus no
C 566	14.8	82.2	204175	2	AC148618	Rattus no	C 639	14.8	82.2	255447	2	AC128889	AC128889	Rattus no
C 567	14.8	82.2	204952	2	AC139999	Rattus no	C 640	14.8	82.2	259354	2	BX364422	BX364422	Danio rer
C 568	14.8	82.2	205893	2	BS000126	Par trogl	C 641	14.8	82.2	259982	2	AL627281	AL627281	Salmonell
C 569	14.8	82.2	207347	9	AC133213	Mus muscu	C 642	14.8	82.2	265050	1	AC126576	AC126576	Rattus no
C 570	14.8	82.2	208370	2	AC133405	Rattus no	C 643	14.8	82.2	265341	3	AE003634	AE003634	Drosophila
C 571	14.8	82.2	208406	2	BX501887	Danio rer	C 644	14.8	82.2	265342	2	AC099138	AC099138	Rattus no
C 572	14.8	82.2	209405	2	AC123601	Mus muscu	C 645	14.8	82.2	266132	2	AC118769	AC118769	Rattus no
C 573	14.8	82.2	209660	2	AC087143	Mus muscu	C 646	14.8	82.2	266374	3	CEY105E8A	CEY105E8A	Rattus no
C 574	14.8	82.2	209873	1	AL414146	Yersinia	C 647	14.8	82.2	271178	3	AC103183	AC103183	Rattus no
C 575	14.8	82.2	210050	2	AC010732	Homo sapi	C 648	14.8	82.2	277607	2	AE003556	AE003556	Drosophila
C 576	14.8	82.2	210164	9	AP001271	Homo sapi	C 649	14.8	82.2	280895	2	AE003485	AE003485	Drosophila
C 577	14.8	82.2	210402	5	BX004964	Zebrafish	C 650	14.8	82.2	283075	3	CH391920	CH391920	Danio rer
C 578	14.8	82.2	211018	5	AC118425	Rattus no	C 651	14.8	82.2	283506	2	SMES591285	SMES591285	Salmonell
C 579	14.8	82.2	211411	1	BS000001	Bacillus su	C 652	14.8	82.2	285806	2	AC116074	AC116074	Rattus no
C 580	14.8	82.2	213080	2	AC126651	Mus muscu	C 653	14.8	82.2	290285	2	AE011733	AE011733	Yersinia
C 581	14.8	82.2	215813	2	AC021477	Rattus no	C 654	14.8	82.2	290874	2	AC137200	AC137200	Rattus no
C 582	14.8	82.2	216120	2	AC111957	Rattus no	C 655	14.8	82.2	291370	2	AC115411	AC115411	Rattus no
C 583	14.8	82.2	216531	2	AC120575	Rattus no	C 656	14.8	82.2	291601	2	AE011857	AE011857	Rattus no
C 584	14.8	82.2	217158	2	AP004168	Oryza sat	C 657	14.8	82.2	291817	1	AE011857	AE011857	Rattus no
C 585	14.8	82.2	217205	8	AC111854	Rattus no	C 658	14.8	82.2	296282	2	AE011857	AE011857	Rattus no
C 586	14.8	82.2	218474	2	CR388047	Danio rer	C 659	14.8	82.2	300029	8	AE0117052	AE0117052	Rattus no
C 587	14.8	82.2	218689	2	AC124666	Mus muscu	C 660	14.8	82.2	300518	2	AE016847	AE016847	Rattus no
C 588	14.8	82.2	219558	2	AC148935	Par trogl	C 661	14.8	82.2	301062	1	AE0117231	AE0117231	Rattus no
C 589	14.8	82.2	220007	2	AL734404	Mouse DNA	C 662	14.8	82.2	301068	1	AE0112393	AE0112393	Mycobacte
C 590	14.8	82.2	220275	10	AL734503	Rattus no	C 663	14.8	82.2	301443	1	BX294143	BX294143	Pirellula
C 591	14.8	82.2	220467	10	AC102504	Mus muscu	C 664	14.8	82.2	303250	1	AE003433	AE003433	Drosophila
C 592	14.8	82.2	221262	2	AC133025	Rattus no	C 665	14.8	82.2	303602	2	AE003509	AE003509	Drosophila
C 593	14.8	82.2	221820	2	AC096427	Rattus no	C 666	14.8	82.2	304390	3	AE003509	AE003509	Drosophila
C 594	14.8	82.2	222469	2	AC101962	Mus muscu	C 667	14.8	82.2	305109	2	AE003509	AE003509	Drosophila
C 595	14.8	82.2	222810	2	AC101962	Mus muscu	C 668	14.8	82.2	305518	2	AE016861	AE016861	Pseudomon
C 596	14.8	82.2	223066	10	AC1017740	Mus muscu	C 669	14.8	82.2	308050	1	AE016861	AE016861	Pseudomon
C 597	14.8	82.2	223761	2	AC1217339	Mus muscu	C 670	14.8	82.2	310029	1	AE016861	AE016861	Pseudomon
C 598	14.8	82.2	223764	2	AC1031366	Rattus no	C 671	14.8	82.2	310581	1	AE016869	AE016869	Pseudomon
C 599	14.8	82.2	223972	2	AC027478	Homo sapi	C 672	14.8	82.2	310967	1	AE016872	AE016872	Pseudomon
C 600	14.8	82.2	224009	2	AC031058	Mus muscu	C 673	14.8	82.2	311963	1	AE003418	AE003418	Drosophila
C 601	14.8	82.2	224915	2	AC130243	Rattus no	C 674	14.8	82.2	314432	3	AE003537	AE003537	Rattus no
C 602	14.8	82.2	225557	2	AC099213	Rattus no	C 675	14.8	82.2	316230	3	AE003666	AE003666	Drosophila
C 603	14.8	82.2	225790	2	AC099213	Rattus no	C 676	14.8	82.2	318278	3	AE003666	AE003666	Drosophila

677	14.8	82.2	322833	2	AC131863	AC131863	Rattus no	750	14.4	80.0	1976	8	SCA304452	AJ034452	Solidago
678	14.8	82.2	326728	2	AC110313	AC110313	Rattus no	751	14.4	80.0	2012	8	AY313607	AY313607	Homo sapi
679	14.8	82.2	334028	1	AC116537	AC116537	Drosophila	752	14.4	80.0	2050	6	CO727169	CO727169	Sequence
680	14.8	82.2	335050	2	AJ141450	AJ141450	Yersinia	753	14.4	80.0	2124	8	AK120224	AK120224	Oryza sat
681	14.8	82.2	340000	9	AP001730	AP001730	Homo sapi	754	14.4	80.0	2129	8	BT012895	BT012895	Lycopodi
682	14.8	82.2	341023	2	AC128351	AC128351	Rattus no	755	14.4	80.0	2147	1	SFL288905	SFL288905	Shigella
683	14.8	82.2	348764	1	EX569689	EX569689	Myoblast	756	14.4	80.0	2197	10	BC021771	BC021771	Mus mus
684	14.8	82.2	348950	1	EX569694	EX569694	Myoblast	757	14.4	80.0	2234	3	AK114461	AK114461	Clona int
685	14.8	82.2	349122	1	EX569699	EX569699	Myoblast	758	14.4	80.0	2269	6	AX747402	AX747402	Sequence
686	14.8	82.2	349122	1	EX569699	EX569699	Myoblast	759	14.4	80.0	2269	6	AK092193	AK092193	Homo sapi
687	14.8	82.2	349122	1	EX569699	EX569699	Myoblast	760	14.4	80.0	2290	6	AX174729	AX174729	Sequence
688	14.8	82.2	349122	1	EX569699	EX569699	Myoblast	761	14.4	80.0	2313	6	AX38692	AX38692	Sequence
689	14.8	82.2	349122	1	EX569699	EX569699	Myoblast	762	14.4	80.0	2335	1	AB031222	AB031222	Sequence
690	14.4	80.0	1174	6	CO806406	CO806406	Sequence	763	14.4	80.0	2537	9	AK097988	AK097988	Homo sapi
691	14.4	80.0	190	9	AB051345	AB051345	Homo sapi	764	14.4	80.0	2703	1	ECOMDB	ECOMDB	Sequence
692	14.4	80.0	293	11	G73536	G73536	Dmpas Mice	765	14.4	80.0	2725	8	AY004212	AY004212	Sequence
693	14.4	80.0	369	6	AX184633	AX184633	Sequence	766	14.4	80.0	2823	8	AP053104	AP053104	Nicotiana
694	14.4	80.0	390	6	CO749899	CO749899	Sequence	767	14.4	80.0	2853	8	AX286518	AX286518	Sequence
695	14.4	80.0	439	6	AX187497	AX187497	Sequence	768	14.4	80.0	2982	8	AK111817	AK111817	Sequence
696	14.4	80.0	468	8	AY491495	AY491495	Venturia	769	14.4	80.0	3001	9	BC030508	BC030508	Sequence
697	14.4	80.0	469	8	CPA431453	CPA431453	Canis fam	770	14.4	80.0	3020	9	BC030508	BC030508	Sequence
698	14.4	80.0	510	6	G89316	G89316	S208P613RD	771	14.4	80.0	3120	1	BACAP	BACAP	Sequence
699	14.4	80.0	556	6	AX186296	AX186296	Sequence	772	14.4	80.0	3148	2	AC019363	AC019363	Sequence
700	14.4	80.0	556	6	CO725781	CO725781	Sequence	773	14.4	80.0	3296	6	BD175474	BD175474	Sequence
701	14.4	80.0	591	6	BD187016	BD187016	ES cell a	774	14.4	80.0	3296	6	AR410854	AR410854	Sequence
702	14.4	80.0	591	6	BD187024	BD187024	ES cell a	775	14.4	80.0	3296	6	AR410854	AR410854	Sequence
703	14.4	80.0	591	10	WMPH34MPA	WMPH34MPA	M. musculus	776	14.4	80.0	3296	6	AR410854	AR410854	Sequence
704	14.4	80.0	593	10	AF450349	AF450349	M. musculus	777	14.4	80.0	3296	6	AR410854	AR410854	Sequence
705	14.4	80.0	595	11	G61734	G61734	SHGC-87084	778	14.4	80.0	3296	6	AR410854	AR410854	Sequence
706	14.4	80.0	642	12	AY657582	AY657582	Synthetic	779	14.4	80.0	3296	6	AX464236	AX464236	Sequence
707	14.4	80.0	648	6	AX617802	AX617802	Sequence	780	14.4	80.0	3296	6	AX697719	AX697719	Sequence
708	14.4	80.0	684	6	AX412600	AX412600	Sequence	781	14.4	80.0	3296	6	BD075623	BD075623	Sequence
709	14.4	80.0	684	6	AX506268	AX506268	Sequence	782	14.4	80.0	3296	6	BD172483	BD172483	Sequence
710	14.4	80.0	684	6	AX651867	AX651867	Sequence	783	14.4	80.0	3296	6	BD172802	BD172802	Sequence
711	14.4	80.0	684	6	BT014828	BT014828	Arabidops	784	14.4	80.0	3296	6	BD173121	BD173121	Sequence
712	14.4	80.0	706	1	AY269009	AY269009	Lactobaci	785	14.4	80.0	3296	9	AY358294	AY358294	Sequence
713	14.4	80.0	740	6	BD220574	BD220574	Human gen	786	14.4	80.0	3389	8	BT002044	BT002044	Homo sapi
714	14.4	80.0	780	8	CNS01993	CNS01993	Sequence	787	14.4	80.0	3389	8	PSY250467	PSY250467	Pinus syl
715	14.4	80.0	783	1	AF164433	AF164433	Bcl-2	788	14.4	80.0	4074	6	CO806432	CO806432	Sequence
716	14.4	80.0	783	6	AR352968	AR352968	Myoblast	789	14.4	80.0	4211	8	KLCP1	KLCP1	Sequence
717	14.4	80.0	805	5	EX931255	EX931255	Gallus ga	790	14.4	80.0	4338	6	CO815989	CO815989	Sequence
718	14.4	80.0	808	6	AX188261	AX188261	Sequence	791	14.4	80.0	4488	6	AR541110	AR541110	Sequence
719	14.4	80.0	812	8	BT012597	BT012597	Arabidops	792	14.4	80.0	4545	2	AC017672	AC017672	Sequence
720	14.4	80.0	819	6	AX620256	AX620256	Sequence	793	14.4	80.0	4940	6	AY2451	AY2451	Sequence
721	14.4	80.0	825	5	AX355873	AX355873	Canis fam	794	14.4	80.0	4940	12	CVE131294	CVE131294	Sequence
722	14.4	80.0	870	5	EX929405	EX929405	Gallus ga	795	14.4	80.0	5239	8	OSAS35058	OSAS35058	Sequence
723	14.4	80.0	926	11	CNS06100W	CNS06100W	Sequence	796	14.4	80.0	5239	2	AC017408	AC017408	Sequence
724	14.4	80.0	951	8	AK104036	AK104036	Oryza sat	797	14.4	80.0	6778	6	CO574688	CO574688	Sequence
725	14.4	80.0	951	8	AK104036	AK104036	Oryza sat	798	14.4	80.0	7401	6	CO598244	CO598244	Sequence
726	14.4	80.0	1068	3	AY008134	AY008134	Caenorhab	799	14.4	80.0	7597	2	AC017434	AC017434	Sequence
727	14.4	80.0	1109	8	AY008134	AY008134	Caenorhab	800	14.4	80.0	7597	2	AC017434	AC017434	Sequence
728	14.4	80.0	1115	8	AK060908	AK060908	Arabidops	801	14.4	80.0	7930	9	HSB08084	HSB08084	Sequence
729	14.4	80.0	1159	6	BD269309	BD269309	Oryza sat	802	14.4	80.0	10431	1	AE002318	AE002318	Homo sapi
730	14.4	80.0	1200	6	BD269309	BD269309	Oryza sat	803	14.4	80.0	11037	2	AE006359	AE006359	Sequence
731	14.4	80.0	1397	5	EX931255	EX931255	Gallus ga	804	14.4	80.0	11067	2	AC014725	AC014725	Sequence
732	14.4	80.0	1413	6	AX838634	AX838634	Sequence	805	14.4	80.0	11905	1	AE006347	AE006347	Sequence
733	14.4	80.0	1413	6	AX838634	AX838634	Sequence	806	14.4	80.0	12464	3	CEC44E1	CEC44E1	Sequence
734	14.4	80.0	1462	6	AR354653	AR354653	Sequence	807	14.4	80.0	12568	3	CEC44E1	CEC44E1	Sequence
735	14.4	80.0	1474	6	AR354653	AR354653	Sequence	808	14.4	80.0	14456	1	AY513488	AY513488	Sequence
736	14.4	80.0	1521	3	AF479563	AF479563	Arabidops	809	14.4	80.0	14787	1	AY513488	AY513488	Sequence
737	14.4	80.0	1640	5	BC078206	BC078206	Danio rer	810	14.4	80.0	14810	2	AC014701	AC014701	Sequence
738	14.4	80.0	1722	5	AF315717	AF315717	Magnapor	811	14.4	80.0	15702	14	AY048850	AY048850	Sequence
739	14.4	80.0	1722	5	AF315717	AF315717	Magnapor	812	14.4	80.0	16341	1	AY048850	AY048850	Sequence
740	14.4	80.0	1816	6	AR339025	AR339025	Sequence	813	14.4	80.0	18211	12	AF327712	AF327712	Cloning v
741	14.4	80.0	1823	6	AR339025	AR339025	Sequence	814	14.4	80.0	20035	2	SPNEU1915	SPNEU1915	Sequence
742	14.4	80.0	1878	8	AX351727	AX351727	Sequence	815	14.4	80.0	21159	2	AC020461	AC020461	Sequence
743	14.4	80.0	1912	8	AY337665	AY337665	Arabidops	816	14.4	80.0	21583	1	AE008838	AE008838	Salmonell
744	14.4	80.0	1914	8	AY337665	AY337665	Arabidops	817	14.4	80.0	24091	9	AC008010	AC008010	Homo sapi
745	14.4	80.0	1931	8	AB032073	AB032073	Nicotiana	818	14.4	80.0	27608	12	AF184978	AF184978	Binary ve
746	14.4	80.0	1942	8	AY397644	AY397644	Solidago	819	14.4	80.0	27719	9	AC012623	AC012623	Homo sapi
747	14.4	80.0	1950	8	E27176	E27176	Novel INPS	820	14.4	80.0	28582	9	AC104822	AC104822	Homo sapi
748	14.4	80.0	1953	8	AF583447	AF583447	Solidago	821	14.4	80.0					
749	14.4	80.0	1959	8	AB009881	AB009881	Nicotiana	822	14.4	80.0					

[illegible]

C 969	14.4	80.0	131152	10	AL929068	AL929068 Mouse DNA
C 970	14.4	80.0	131441	8	AL954871	Oryza sat
C 971	14.4	80.0	131692	8	AC006423	Arabidops
C 972	14.4	80.0	132728	9	AC024121	Human DNA s
C 973	14.4	80.0	133568	9	HS16915	Human DNA s
C 974	14.4	80.0	133615	2	AC146749	Medicago
C 975	14.4	80.0	134355	8	AC124952	Medicago
C 976	14.4	80.0	134532	9	AP006442	Homo sapi
C 977	14.4	80.0	134712	2	AF003891	Oryza sat
C 978	14.4	80.0	134741	10	AL626805	Mouse DNA
C 979	14.4	80.0	134741	10	AL683819	Mouse DNA
C 980	14.4	80.0	134938	2	AC137759	Oryza sat
C 981	14.4	80.0	135105	8	CNS08CAB	Oryza sat
C 982	14.4	80.0	135660	2	AC103927	Mus muscu
C 983	14.4	80.0	135820	2	AL513497	Human DNA
C 984	14.4	80.0	136657	10	AC121572	Mus muscu
C 985	14.4	80.0	137122	10	AC123037	Mus muscu
C 986	14.4	80.0	137674	2	AC117697	Mus muscu
C 987	14.4	80.0	138020	9	AC108043	Homo sapi
C 988	14.4	80.0	138231	9	AC093558	Homo sapi
C 989	14.4	80.0	138264	9	AC008722	Homo sapi
C 990	14.4	80.0	138633	2	AL356360	Homo sapi
C 991	14.4	80.0	139084	5	AL353332	Zebrafish
C 992	14.4	80.0	139408	2	AC079565	Mus muscu
C 993	14.4	80.0	140750	2	AC131239	Medicago
C 994	14.4	80.0	140944	10	AL450331	Mouse DNA
C 995	14.4	80.0	141458	2	AC142044	Rattus no
C 996	14.4	80.0	142273	9	AL391994	Human DNA
C 997	14.4	80.0	142475	8	AP003627	Oryza sat
C 998	14.4	80.0	143206	2	AC141776	Apis mell
C 999	14.4	80.0	143575	2	AC092370	Homo sapi
C1000	14.4	80.0	143765	8	AP004623	Oryza sat

ALIGNMENTS

RESULT 1
CR382126 06/c
WPCOMMENT
Sequence split into 26 fragments LOCUS CR382126 Accession CR382126

Fragment Name	Begin	End
CR382126_00	1	110000
CR382126_01	100001	210000
CR382126_02	200001	310000
CR382126_03	300001	410000
CR382126_04	400001	510000
CR382126_05	500001	610000
CR382126_06	600001	710000
CR382126_07	700001	810000
CR382126_08	800001	910000
CR382126_09	900001	1010000
CR382126_10	1000001	1100000
CR382126_11	1100001	1210000
CR382126_12	1200001	1310000
CR382126_13	1300001	1410000
CR382126_14	1400001	1510000
CR382126_15	1500001	1610000
CR382126_16	1600001	1710000
CR382126_17	1700001	1810000
CR382126_18	1800001	1910000
CR382126_19	1900001	2010000
CR382126_20	2000001	2110000
CR382126_21	2100001	2210000
CR382126_22	2200001	2310000
CR382126_23	2300001	2410000
CR382126_24	2400001	2510000
CR382126_25	2500001	2602197

Continuation 77 of 26) of CR382126 Kluyveromyces fragilis strata
Query Match 94.4%; Score 17; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACCAACAGATTGG 17
DB 81584 GAGAACCAACAGATTGG 81568
RESULT 2
AE006452/c 10610 bp DNA linear BCT 02-JUN-2004
LOCUS Lactococcus lactis subsp. lactis 11403 section 214 of 218 of the
DEFINITION complete genome.
ACCESSION AE006452 AE005176
VERSION AE006452.1 GI:12725297
KEYWORDS
SOURCE
ORGANISM Lactococcus lactis subsp. lactis 11403
Lactococcus lactis subsp. lactis 11403
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 10610)
Mauger, S., Jallion, O., Malarme, K.,
Bojotin, A., Wincker, P., Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
The complete genome sequence of the lactic acid bacterium
Lactococcus lactis ssp. lactis 11403
Lactococcus Res. 11 (5), 731-753 (2001)
JOURNAL
MEDLINE
PUBMED 21235186
REFERENCE 2 (bases 1 to 10610)
Mauger, S., Jallion, O., Malarme, K.,
Bojotin, A., Wincker, P., Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
Direct Submission
Submitted (09-JUN-2001) INRA, Genetique Microbienne, Domaine de
Vilvert, Jouy en Josas 93352, France
LOCATION/Qualifiers
1. 10610
/organism="Lactococcus lactis subsp. lactis 11403"
/mol_type="genomic DNA"
/strain="11403"
/sub_species="lactis"
/db_xref="taxon:272623"
complement (83..610)
/gene="mreD"
/note="synonym: L98132"
complement (83..610)
/gene="mreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 Murein sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=11
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIITAIITIAISAFISAKNFKNRNSANTQTI
NDGTSFVDRVLAIPHVDKANEENIMDTYQOQOSLKTQLAKSKDDDKLSGLESE

gene
CDS
/gene="mreD"
/note="synonym: L98132"
complement (83..610)
/gene="mreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 Murein sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=11
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIITAIITIAISAFISAKNFKNRNSANTQTI
NDGTSFVDRVLAIPHVDKANEENIMDTYQOQOSLKTQLAKSKDDDKLSGLESE

gene
CDS
/gene="mreD"
/note="synonym: L98132"
complement (700..1575)
/gene="mreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 Murein sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=11
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIITAIITIAISAFISAKNFKNRNSANTQTI
NDGTSFVDRVLAIPHVDKANEENIMDTYQOQOSLKTQLAKSKDDDKLSGLESE

gene
CDS
/gene="mreD"
/note="synonym: L98132"
complement (700..1575)
/gene="mreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 Murein sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=11
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIITAIITIAISAFISAKNFKNRNSANTQTI
NDGTSFVDRVLAIPHVDKANEENIMDTYQOQOSLKTQLAKSKDDDKLSGLESE

gene
CDS
1868..2308
/gene="yxbC"
/note="synonym: L99869"
1868..2308
/gene="yxbC"
/note="EVIDENCE BY HOMOLOGOY BIO15.01 HYPOTHETICAL."
General"
/codon_start=1
/transl_table=1
/product="HYPOTHETICAL PROTEIN"
/protein_id="AAK06330.1"
/db_xref="GI:12725300"
/translation="NMHI PVFENKNEKNQARVGVYSIVARENKEICLVQAPNGAYF
LPGBIEAGEDQEKALHRELMELGATLIEFGADDEYFSSHRDKYYNPAIYIN
TKNVFDSKPLEDFNHSIMDPETFAISKLKSGSHKMGVQEWLKK"
complement(2336..3010)
/gene="piA"
/note="synonym: L0045"
complement(2336..3010)
/gene="piA"
/note="EVIDENCE BY HOMOLOGOY BIO06.10 Pentose phosphate
pathway"
/codon_start=1
/transl_table=1
/product="ribose 5-phosphate isomerase A (EC 5.3.1.6)"
/protein_id="AAK06331.1"
/db_xref="GI:12725301"
/translation="MDNLKQVGIKAEFVKSQVGVGLGISTAYFVEELGRIAE
OLEITGVTSVNTSQARALGIPLASIDRVYVDTVDGADLSDSLNGKGGALL
MEKIVATSKYIIVDESKLSENGSFKIPEVYIPYGOQVFKFPAAGVAPRPLN
BENRLITDMHFIIDHISQIKPEKKAEBLDLMGVVHGLFNNVKKVIVAGNEG
VRIINK"
complement(3094..3780)
/gene="rcfB"
/note="synonym: L0245"
complement(3094..3780)
/gene="rcfB"
/note="EVIDENCE BY HOMOLOGOY BIO09.01 REGULATORY FUNCTIONS.
General. BELONGS TO THE CRP/FNR FAMILY OF TRANSCRIPTIONAL
REGULATORS"
/codon_start=1
/transl_table=1
/product="transcriptional regulator"
/protein_id="AAK06332.1"
/db_xref="GI:12725302"
/translation="WSPFKMEEMANVNIYLLINTEOKKVVTRKRSYIMYGIES
NYIKYKGVAKISITLDRGEFNIAVYADPVSILBEKONGISALFNVVRESPPA
NFKYISRDPMWVREDLNUFRVVDPEFKRLANLBILOKTINGKRGVACASLH
IDDFIRKDDGLIDFTVNEBIAFGCISTRSVNRILHLDKDEKVIDIVDKIMTY
NFOYLEEYIS"
complement(3863..5083)
/gene="yxbD"
/note="synonym: L101912"
complement(3863..5083)
/gene="yxbD"
/note="EVIDENCE BY HOMOLOGOY BIO13.08 TRANSPORT AND BINDING
PROTEINS. Multidrug resistance, similar to quinolone
resistance protein Nora from difference species"
/codon_start=1
/transl_table=1
/product="transporter"
/protein_id="AAK06333.1"
/db_xref="GI:12725303"
/translation="MKKESIRPTPVNTLISFLIVFYVLTSGISTVALQLOHST
VSLSSVFIYVIGALIGRWGTINIRIGKRLIYIGLRLVLTFRGYITNPLF
LIRVIQAGFGIATAGVAVPARREGIGYALSVTLAAVGPALISITVYS
LFGSLIGLIGLIVITLIFVYVKEFSABEAYALAHDPKVEYIIEKSLPISI
IAFLAGFISALITGMSFSDIKIPLAGSLFTMYAVLIFVSRPTGRLFDTKGDMW
INPTFIIPAVAMLVGLAGFSPAIQVIVLLIAGAVGLGVAAPFGQALAIRDS

gene
CDS
complement(5146..5583)
/gene="yxbE"
/note="synonym: L103195"
complement(5146..5583)
/gene="yxbE"
/note="EVIDENCE BY HOMOLOGOY BIO15.02 HYPOTHETICAL.
conserved, there are six paralogs in L.lactis"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAK06334.1"
/db_xref="GI:12725304"
/translation="MEQKISRIIVGMDGSAESLKAPDEAVRVAKRNDAILANVEX
RSFLAVNYDTIAEGBHNGTKLLQOPFAEALDEGLTNRITLEYGSPRVAAATKIP
KQENVDLIVIGATGISYIERVYGVSVASYVNHATCNVILVRK"
complement(5603..5806)
/gene="yxbF"
/note="synonym: L103652"
complement(5603..5806)
/gene="yxbF"
/note="EVIDENCE BY HOMOLOGOY BIO15.01 HYPOTHETICAL.
General"
/codon_start=1
/transl_table=1
/product="HYPOTHETICAL PROTEIN"
/protein_id="AAK06335.1"
/db_xref="GI:12725305"
/translation="MKKESIRLSYEENSTPNVYGTKRASVYQDFVNDDEALFEFVN
HHGFKTBDGKFAFVERNISATG"
complement(6066..10376)
/gene="yxbA"
/note="synonym: L104115"
complement(6066..10376)
/gene="yxbA"
/note="EVIDENCE BY HOMOLOGOY BIO15.02 HYPOTHETICAL.
conserved, 30% identical to activator of
(R)-2-hydroxyglutaryl-CoA dehydratase of Acidimicrococcus
fermentans."
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAK06336.1"
/db_xref="GI:12725306"
/translation="WTOIKRTGIDVSGTIVKLVVFDENFKLISRIERHSDVKATI
KVFDFIEYEDLNLBELAINGSGMGAELVIVFVQEVIASSITIEKFIPTDVMIE
LGEDAKVTFPDNGTOELRNNGTCAGGTGAIFQVMAALLVDANGVIELKKNKXP
IASRGVAKTQVQPLNDEGRREDIASIFQAVVNTTIGLGSGRKIKNVAFILGP
LFPSELEKREIFETLDKEKNTIPENPOLFVANGALYEETSOQXTDILHLEND
SDSLIPDMMKLFESQALDESPRTAKXTANYNKLSHEHGAALFGLIDGISTTKV
LIDEGNTLVHGSNGGPELESYINILKOVYAKLPADVYIASVTVGSEHILKAGL
HVDGEVETMHHYRAADYFNPQVDFIIDIGQDKAMRIDGALSSIQLEKASSGSG
SFIETPASTLHDVREFSQALLAEHVNIGSRCTVPMKRVKQVQEGATVDISAG
LSYIKNALYKVIKVRPELIKRTVQGTGFYENALVSAFVKISEREVVRSLIGL
MGVAGCALISLNDQANKSEHILKLDLEFTTHKEVNLGLENKCMKLTVPENGN
KEVTCGECERAEKATYKVAKKDKKNLVADYKKLFRHSHSKKQTEGEGITPRV
LMYENYVLMHTMLTDLGFRVLSFRSDEKLFEBGIEITIPSDYCVYAKSHGHIAL
IKOQIPNTPVPSVFEQEBOKNANGNCIIVGSPYV.LKNMIDEBEGOVNLHPI
NLANDGEVAVVNHKLTROGISVLTETVQVAVOHGPFEMKPFEDLRLKREELMOIN
LNNEKATYLRPPIHLPETNHGADIITDGBGHVLTREDSISHLABGSRVYNQVY
HSRLPAANVVCNKNLLELVQNSFGGLDAVTTDVEELMRGNKLYTLKIDBSN

Query Match 91.1% Score 16.4; DB 1; Length 10610;
Best Local Similarity 94.4%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 5871 GGAACCAACAGATTCCC 5854
OY 1 GGAACCAACAGATTCCC 18
RESULT 3

```

MLCB1788
LOCUS       MLCB1788             39228 bp    DNA             linear      BCT 27-AUG-1999
DEFINITION  Mycobacterium leprae cosmid B1788.
ACCESSION   AL008609
VERSION     AL008609.1 GI:25783368
KEYWORDS    43L; 6-phosphogluconate dehydrogenase; ABC-transporter; antigen
            43L; cytochrome P450; dehydrogenase; gcwH; gcwJ; glcB; glycine
            cleavage system h protein; glycine dehydrogenase (decarboxylating);
            gnd; gnaB; inosine-5'-monophosphate dehydrogenase; LEPRP; malate
            synthase G; oxidoreductase; pgsA; phosphatidyltransferase;
            preprotein translocase secA subunit; pseudogene; secA;
            TPP-regulating enzyme.
SOURCE      Mycobacterium leprae
ORGANISM    Bacteria; Actinobacteria; Actinobacteriales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE   1 (bases 1 to 39228)
AUTHORS     Eigmeier, K., Honore, N., Woods, S.A., Caudron, B. and Cole, S.T.
TITLE       Use of an ordered cosmid library to deduce the genomic organization
            of Mycobacterium leprae
JOURNAL     Mol. Microbiol. 7 (2), 197-206 (1993)
MEDLINE     8446027
REFERENCE   2 (bases 1 to 39228)
AUTHORS     Skelton, J. and Churcher, C.M.
TITLE       Unpublished
JOURNAL     3 (bases 1 to 39228)
AUTHORS     Parkhill, J., Barrell, B.G. and Rajandream, M.A.
TITLE       Direct Submission
JOURNAL     Submitted (21-Oct-1997) Mycobacterium leprae sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Dr.
            Stewart T. Cole, [3] Unite de Genetique Moleculaire Bacterienne,
            Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15,
            France Requests for cosmids should be sent to Karin Eigmeier
            (k.eigmeier@pasteur.fr)
COMMENT     Notes:
            The Sanger Centre is funded to complete the sequence of M. leprae
            by the Helser Program for Research in Leprosy and Tuberculosis of
            The New York Community Trust.
            Work in Paris is supported by the Helser Trust, the Association
            Française Raoul Follereau and the Groupement de Recherches et des
            Etudes des Genomes (GEP-GRG).
            Details of M. leprae sequencing at the Sanger Centre are available
            on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/)
            CDS are numbered using the following system eg MLCB33.01c. ML (M.
            leprae), CB33 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous. The length
            in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon. All CDS
            over 100 codons have been analysed. Gene prediction is based on
            positional base preference in codons especially where there is an
            increase in the observed/expected third position G + C. CAUTION:
            We may not have predicted the correct initiation codon. Where
            possible we choose an initiation codon (atg, gtc, or ttg) which is
            preceded by an upstream ribosome binding site sequence (optimally is
            5-13bp before the initiation codon). If this cannot be identified
            we choose the most upstream initiation codon.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone. It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions. Cosmid B1788 is
            overlapped at the 5' end by EM_BA:MLB38COS.
FEATURES             source
            1..39228
            /organism="Mycobacterium leprae"
            /mol_type="genomic DNA"
            misc_feature
            1..2183
            /note="overlap with EM_BA:MLB38COS L01095 M. leprae cosmid
            B38 from: 34883 to: 37065"
            complement(1..256)
            /gene="43L"
            /gene="43L"
            complement(<1..256)
            /gene="43L"
            /note="MLCB1788.01c. 43L antigen 43L partial CDS, len:
            >85 aa; almost identical to M43L_MYCBL P46842 antigen 43L
            precursor (97.6% identity in 85 aa overlap), also highly
            similar to M. tuberculosis TR:050906 (EMBL:X80268), apa,
            45/47 gene precursor (325 aa), fasta scores: opt: 246
            z-score: 289.6 E() : 5.6e-09, 52.1% identity in 96 aa
            overlap. Contains N-terminal signal sequence"
            /codon_start=1
            /transl_table=1
            /product="antigen 43L"
            /protein_id="CAA15433.1"
            /db_xref="GI:25783369"
            /translation="MNQVLDSTHRKGMALIAIAYVASAFITMPLPAANADPAPL
            PSTATAPSPAQELITPLPGAPVSEHQPDPRNAPSLDPN"
            35
            /note="conflict: T is C in EMBL:MLB38COS"
            complement(265..270)
            /note="possible RBS upstream of 43L"
            complement(482..787)
            /note="MLCB1788.02c. ABC-transporter pseudogene fragment,
            len: 101 aa; highly similar to part of TR:P95155 005126,
            MTCY359.14c. modC, ATP-binding subunit of ABC-transport
            system (369 aa), (75.0% identity in 56 aa overlap)"
            /pseudo
            /codon_start=1
            /transl_table=1
            /db_xref="PSEUDO:CAA15434.1"
            complement(863..1033)
            /note="MLCB1788.03c. ABC-transporter pseudogene fragment,
            len: 56 aa; highly similar to part of TR:P95155 005126,
            MTCY359.14c. modC, ATP-binding subunit of ABC-transport
            system (369 aa), (75.0% identity in 56 aa overlap)"
            /pseudo
            /codon_start=1
            /transl_table=1
            /db_xref="PSEUDO:CAA15435.1"
            complement(1193..1321)
            /note="MLCB1788.04c. ABC-transporter pseudogene fragment,
            len: 42 aa; highly similar to part of TR:P95155 005126,
            MTCY359.14c. modC, ATP-binding subunit of ABC-transport
            system (369 aa), (81.0% identity in 42 aa overlap)"
            /pseudo
            /codon_start=1
            /transl_table=1
            /db_xref="PSEUDO:CAA15436.1"
            complement(1315..1467)
            /note="MLCB1788.05c. ABC-transporter pseudogene fragment,
            len: 50 aa; highly similar to part of TR:P95155 005126,
            MTCY359.14c. modC, ATP-binding subunit of ABC-transport
            system (369 aa), (70.0% identity in 50 aa overlap).
            Contains P950017 ATP/GTP-binding site motif A (P-loop)"
            /pseudo
            /codon_start=1
            /transl_table=1
            /db_xref="PSEUDO:CAA15437.1"
            complement(1354..1377)
            /note="P950017 ATP/GTP-binding site motif A (P-loop)"
            /note="MLCB1788.06c. ABC-transporter pseudogene fragment,
            len: 64 aa; highly similar to part of TR:P95155 008061,
            MTCY359.15c. modB, membrane spanning subunit of
            ABC-transport system (264 aa), (67.2% identity in 64 aa
            overlap)"

```

```

CDS
/pseudo
/codon_start=1
/transl_table=11
/note="MNCB1788.07c, ABC-transporter pseudogene fragment,
len: 89 aa; highly similar to part of TR:P95156 O08061,
MTCY359.15c, modB, membrane spanning subunit of
ABC-transport system (264 aa), (60.7% identity in 89 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/db_xref="PSEUDO:CA15439.1"
/note="MNCB1788.08c, ABC-transporter pseudogene fragment,
len: 34 aa; highly similar to part of TR:P95156 O08061,
MTCY359.15c, modB, membrane spanning subunit of
ABC-transport system (264 aa), (68.8% identity in 32 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/complement(2418..2804)
/note="MNCB1788.09c, ABC-transporter pseudogene fragment,
len:128 aa; highly similar to part of TR:P95157 O05125,
MTCY359.16c, modA, molybdenum-binding subunit of
ABC-transport system (261 aa), (42.3% identity in 97 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/db_xref="PSEUDO:CA15441.1"
/complement(2801..3058)
/note="MNCB1788.10c, ABC-transport system pseudogene
fragment, len: 85 aa; highly similar to part of TR:P95157
O05125, MTCY359.16c, modA, molybdenum-binding subunit of
ABC-transport system (261 aa), (61.2% identity in 85 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/codon_start=1
/transl_table=11
/note="possible RBS upstream of MNCB1788.11c"
3300..3306
3308..3382
/gene="MNCB1788.11c"
3308..3382
/gene="MNCB1788.11c"
/note="MNCB1788.11c, probable oxidoreductase, len: 224 aa;
similar to eg. FBG_ECOLI_P5716 3-oxoacyl-(acyl-carrier
protein) reductase (244 aa), fasta scores; opt: 157
z-score: 273.8 E(): 4.2e-08, 27.9% identity in 240 aa
overlap. Also highly similar to M. tuberculosis
oxidoreductase MTCY359.17 (80.4% identity in 224 aa
overlap)"
/codon_start=1

Query Match 91.1%; Score 16.4; DB 1; Length 39228;
Best Local Similarity 94.4%; Pred. No. 5,4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GAGAACAAAGATTGCC 18
DB 15222 GAGTACAAAGATTGCC 15239

RESULT 4
AC008047/c 100867 bp DNA linear PLN 28-JUN-2000
LOCUS AC008047
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F2K11 from chromosome
1, complete sequence.
ACCESSION AC008047
VERSION AC008047.3 GI:6623889
KEYWORDS HTG.

```

```

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 100867)
REFERENCE
AUTHORS
Shinn, P., Brooks, S., Buehler, E., Chao, O., Johnson-Hopson, C.,
Khan, S., Kim, C., Altafi, H., Bei, Q., Chin, C., Chiu, J., Choi, E.,
Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T.,
Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, K., Liu, S.,
Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M., Yu, G.,
Fedorisiel, N.A., Theologis, A. and Ecker, J.R.
Genomic sequence for Arabidopsis thaliana BAC F2K11 from chromosome
1
I
Unpublished
2 (bases 1 to 100867)
REFERENCE
AUTHORS
Ecker, J.R.
Direct Submission
Submitted (15-JUN-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 100867)
REFERENCE
AUTHORS
Ecker, J.R.
Direct Submission
Submitted (21-DEC-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 100867)
REFERENCE
AUTHORS
Shinn, P., Brooks, S., Buehler, E., Chao, O., Johnson-Hopson, C.,
Khan, S., Kim, C., Altafi, H., Bei, Q., Chin, C., Chiu, J., Choi, E.,
Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T.,
Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, K., Liu, S.,
Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M., Yu, G., Davis, R.,
Fedorisiel, N., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (23-DEC-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
5 (bases 1 to 100867)
REFERENCE
AUTHORS
Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, O.,
Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, Q.,
Chiu, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N.,
Howing, B., Koo, T., Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, J.,
Liu, S., Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwarz, U., Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M.,
Yu, G., Davis, R., Fedorisiel, N., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
On Dec 21, 1999 this sequence version replaced gi:5870168.
COMMENT
FEATURES
source
1..100867
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clome="F2K11"
join(1694..1814,2131..2282,2395..2493,2874..3083,
3173..3256,3400..3477,3611..3668,4074..4415,4496..4577,
4675..4770,4859..5232,5218..5458,5578..5822,5514..6167,
6274..6437,6543..7129,7218..7437)
/note="similar to kinesin-like protein gi|3392771, similar
to Bets|gb|U60306, gb|A1560436.1, gb|T22003, gb|W43855,
and gb|A585767"
/codon_start=1
/evidence=not experimental
/product="F2K11.1"
/protein_id="AA19694.1"
/db_xref="gi:6633833"
/translation="VSSHLSODANWNGVYVRSDVSVSMLSFDGSEBSREKSDSKGKHOS

```

```
LYEWNETLPYKLPWEASEDELRACLRDGYLCSLLNQLSPGSMGSGFEPASVKI
ERFLTAMDMLPREVSDIQSLKALNKFSDGSDYKNSLARRMSLPDHSNRG
DNRNFDGQSGESGIDMSDAKISDLKSNLRAPRSLEPMDLKLDSMSRNG
HVSNAWASLISLVYIENIRNOADNLKVNKGLRILGSDWNGKRGKTVTSSLE
KTRIEKESSEKDVYRLKREKERSDAPOLKYLKHEHOCLELAKKOKTR
DELEKIKDAELHVVDSRKYELKELCOSKORKEKCECTYONFIDHNSALDELISA
TSLSIKHEVYRORXYFEDNLYGLKGVADAAKHYVLEENRRLNVEOLKGNL
RYCHIRPPLPGQNSRQTTIEYIGTGLVWANKPKQDTRLEFKYKVOQATOG
LNLFPSSVYLAFSS11YDIHLINLISVEEVLDRPLRSLLDGYNCIPAYQOTS
GLTYMSGVLSITSKEDGVNRYALNDLFLTQSRONTYMEGVQWVEIYEGONMG
LVLTGFTVNLMSRIONFLNHTLIGINTALPNGLAVPADAMHCVRSTEDYLEMNG
LMNRVGAATLNERSSRSHCVLSVHNGVDVDEDS1LDSGLHVLVLAGSRVDSERT
GERLKEAOKHINKSLSALGDVYFALAKRPHPRYNSKLTQVLOSGLGQAKTLNFVY
NPDGSYALVTSITLKFARVSGVELGAASKSREGDVRQMEQVSNLKDVIAKDEEL
ONFOYKGNMATSLSKRGSLNRLVGPISPRHS1GASPNARAKASGLTGRGTSVDVN
CEYSKHSIDSGSQSDSRKQMDYHQPSFAGAKGIDPDDEVELVGLDSDSD
RLSDISDCLMGTETDGSISAVELLFETAKPLIEREBAMTSSEKLEKVGK
KTEPDSTRNIPSKIKOTLKPQOTRPSRLS1ATSSSSKXLTNGS1HQLFDLCPRN
OICPLTILIGSNHSLRFL"
complement (1872, .9535, 9606, .97401)
/Note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F2K11.2"
/protein_id="AA19720.1"
/db_xref="GI:6633861"
/translation="MWENGCPNVTYNTLNNSICRGKRVLEALAVDRVESHOPN
AMGSHIKADYVISTALYDRICKGNHIMKNTLTEMHKGIFPNVLTNMCIDSFCH
SGNSADQILRHMIKQINPDIVTFSALINAFKERVSEAEIYEMLRMSI1PFT
ITNSMIDGFCQKORVDKAKMLDSMKSQSPVVTSTLNGCYAKKADVMEIF
CEMHRGRIVANTVTYTLIHGFCQVGLDAADILNEMI1SGVAPDYITTHCLAGC
SKELRAFAILEDLOKSEVCPRLIFLFSICSLIGFGEVCVRIYOVKRIK1SD"
join (11234, .11473, 11567, .11656, 11735, .11815, 11913, .12033,
12547, .12683)
/Note="unknown; similar to ESTs gb|AM032410,
gb|AM053721.1, and gb|A199565.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.3"
/protein_id="AA19695.1"
/db_xref="GI:6633836"
/translation="MCSLSVQFSLIOSPPSPPLANHEPLTSSSTVTPLEKTN
TKOSGIGNLVLRKAVYSSSDSSADSTTPNGTRKSRDILLEYVNVNPEFEM
FVKRAPHVVEAKRQVTNMI1GTLPPQFPAVTVAENIACLMASVLVGNFRAO
YRLLEIQSLEQVALPEPRGITLYALRKLKMSKISSPWKQSKSTVWKQENQIRSMK
SWSI"
complement (14209, .15126)
/Note="similar to receptor-like protein kinase gi|4008012;
similar to ESTs gb|T21898 and gb|R90176"
/codon_start=1
/evidence=not experimental
/product="F2K11.4"
/protein_id="AA19719.1"
/db_xref="GI:6633860"
/translation="WATMTQSLSLFCFSLISLFLSTVLTLSGSDHIMSTFCQFS
DNFTQSYETNRETLLSLRLBSLIGSTSNATAGISPTVGMFLCRGDISTSCSD
CVTATLISRKTQKKA1FYEBECWVYSDSSPFSVLDERP1IRYSLASYANLDR
PQTLSDKMDLILNATSSPSLSTPYFVEDERVSFDIDSMACSPDLDRN
CTTCLAVOEMLECCNOSQMAO1FTPKCLLVEATLALSSPPPPYSPSPSPSP
LPSPLFRKPTQASGSGSSSINVKM"
complement (16131, .16937)
/Note="similar to serine/threonine kinase-like protein
emb|CA18465; similar to EST gb|A1994269.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.5"
/protein_id="AA19718.1"
/db_xref="GI:6633859"
/translation="WNGINTHRAVALFCFSSLRAMQASQML1DCTQPDNVTSS
YLSNRPTVSTIRNNSISGSINATAGSPNTYGMFLCRGDLNNTSCDQVNA1TTL
TKSCFIRASALVINSECIKRVSNVSPFLVEDVSTARFSTGNSLDSGPFQSGTILE
KDALILRASLSSSPVYFVDDQOHVTLDESDVLAHMAVQCSPLDPRNCTVCIRLA
VRLSCCASHAQPARIFYTKCLITYEISALQNVNTSLAVTSESILVY11"
CDS
complement (18736, .19500)
/Note="similar to protein kinase-like protein
emb|CA18472; similar to ESTs gb|A1994269.1 and gb|h36750"
/codon_start=1
/evidence=not experimental
/product="F2K11.6"
/protein_id="AA19717.1"
/db_xref="GI:6633858"
/translation="WRRITDVLSLFCFLFLSLAMQSPGNLTDNQPDPTQSNNTV
LSLRLHSSLSGYFRATIGLSLNTVYGMFLCTIGN1SKTSCSNVASHATLENDKSES
HDSFEMFSECVARYSDNSFSLVDSFATFSYSONDSLSPQFYNQTLPGCLDELIL
KAPSSFSVPVYFVEDKEHVTQVESYDLEMAAQCSPTLDPSSCTVCLGLVEKSEFC
CSQSRWAR1HFPKCLIRYD1SALQCNLSLGLVTKG"
complement (20529, .21542)
/Note="similar to kinase-like protein emb|CA18472;
similar to ESTs gb|A1994269.1, gb|h21898, gb|h36750, and
gb|R90176"
/codon_start=1
/evidence=not experimental
/product="F2K11.7"
/protein_id="AA19716.1"
/db_xref="GI:6633857"
/translation="WRKXTSFTYSSSESLVLPYVLDITLKSTKSSPHTNSLSP
LFRITQNMARIT1TALFFRFFSLSHQMSQPHMIFCSVDSFTQTSYSTNRN
ILITLHSLSSLVHYNATIGLSPDTVGMFLCRGDINTTSCQCVQAA1E1A1NCT
LNKRAFLIYDECMERYSNVSEFSESPVIRVYSLRSPNSRPNQTLSSKDLQ1P
NVPSTLLPYFVEDQERYTQLEGSYDLVSMIQSPDLDPSCITCLRAVATVETCCG
VPSA11FTPKCII1RRTVFLPSPAPSPSS1P1SPSTSPPLSLPQLPPLSQPPP
LS1"
complement (22893, .23747)
/Note="similar to serine/threonine protein kinase
emb|CA18705.1; similar to ESTs gb|R90176 and
gb|A1994269.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.8"
/protein_id="AA19715.1"
/db_xref="GI:6633856"
/translation="WATFIRFTAPLFCFFFLFLSFSHQMSQPHM1FCDRLASNF
TQTSYKXANRETLISLDRSLGTQYNATIGLSPDTVGMFLCRGDLTTSQSDCVK
TATLEITQNNCTSRKTAL1FYEBCMRVASPSFTLDEGDPVAKVATPSTSLF
SSFOOTLSNRYEOL1ILIASKSLASSTPYVXQISVNELEGSYTLDTVACQSPDL
PANCQPLRV1VKNGLSQCCTNARAQY1LCKFLKYTTTLPTSQSSSGSSINVMKG
HNGCPL"
complement (26076, .27047)
/Note="similar to protein kinase-like protein
emb|CA18472; similar to ESTs gb|T21898 and gb|A1994269.1"
/codon_start=1
Query Match 91.1%; Score 16.4; DB 8; Length 100867;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS
1 GAGAACAAAGATTCCG 18
DB 69213 GAGAACAAAGATTCCG 69196
CDS
complement (126076, .27047)
/Note="similar to protein kinase-like protein
emb|CA18472; similar to ESTs gb|T21898 and gb|A1994269.1"
/codon_start=1
Sequence split into 36 fragments LOCUS LMFLCHR36 Accession AL499624
WPCOMMENT
LMFLCHR36_27
LMFLCHR36_00 1 110000
LMFLCHR36_01 100001 210000
LMFLCHR36_02 200001 310000
LMFLCHR36_03 300001 410000
LMFLCHR36_04 400001 510000
LMFLCHR36_05 500001 610000
LMFLCHR36_06 600001 710000
LMFLCHR36_07 700001 810000
LMFLCHR36_08 800001 910000
LMFLCHR36_09 900001 1010000
```


LMFLCHR36_10 1000001 1110000
 LMFLCHR36_11 1100001 1210000
 LMFLCHR36_12 1200001 1310000
 LMFLCHR36_13 1300001 1410000
 LMFLCHR36_14 1400001 1510000
 LMFLCHR36_15 1500001 1610000
 LMFLCHR36_16 1600001 1710000
 LMFLCHR36_17 1700001 1810000
 LMFLCHR36_18 1800001 1910000
 LMFLCHR36_19 1900001 2010000
 LMFLCHR36_20 2000001 2110000
 LMFLCHR36_21 2100001 2210000
 LMFLCHR36_22 2200001 2310000
 LMFLCHR36_23 2300001 2410000
 LMFLCHR36_24 2400001 2510000
 LMFLCHR36_25 2500001 2610000
 LMFLCHR36_26 2600001 2710000
 LMFLCHR36_27 2700001 2810000
 LMFLCHR36_28 2800001 2910000
 LMFLCHR36_29 2900001 3010000
 LMFLCHR36_30 3000001 3110000
 LMFLCHR36_31 3100001 3210000
 LMFLCHR36_32 3200001 3310000
 LMFLCHR36_33 3300001 3410000
 LMFLCHR36_34 3400001 3510000
 LMFLCHR36_35 3500001 3529952

Continuation (28 of 36) of LMFLCHR36 from base 2700001 (AL499624 Leishmania major chromo

Query Match 91.1% Score 16.4; DB 2; Length 110000;
 Best Local Similarity 94.4%; Pred. No. 5.2e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCCG 18
 |||||
 Db 91900 GAGACCAACAGATTCCG 91917

RESULT 6
 AC143118/c 124215 bp DNA linear HTG 09-APR-2003
 LOCUS Macaca mulatta clone CH250-269J3, *** SEQUENCING IN PROGRESS ***.
 DEFINITION AC143118.1 GI:29567757
 ACCESSION AC143118
 VERSION HTG; HTGS_PHASE2; HTGS_PGI
 KEYWORDS Macaca mulatta (rhesus monkey)
 SOURCE Macaca mulatta
 ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
 1 (bases 1 to 124215)
 Csuros,M. and Milosavljevic,A.
 Pooled genomic indexing (PGI): mathematical analysis and experiment
 design
 (in) Guigo,R. and Gusfield,D. (Eds.);
 ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
 Springer (2002)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 124215)
 Milosavljevic,A., Sodergren,E., Csuros,M., Li,B., Jackson,A.R.,
 Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsdicks,S.L.,
 Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbata,J.,
 Benton,J., Blum,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowe,S., Brieva,M., Brown,E., Bryant,M., Bryant,N.P., Buñay,C.,
 Burcher,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Cleveland,C.D.,
 Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleland,R.,
 Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rech-S., Durbin,K.J., Egan,A., Earnhart,C., Edwards,C.C.,
 Elhai,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
 Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Geo,J.,
 Garcia,A., Garner,T., Garza,N., Gill,R., Gotreill,U.H., Guevara,W.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 3 (bases 1 to 124215)
 Worley,K.C.
 Direct Submission
 Submitted (05-APR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 124215)
 Worley,K.C.
 Direct Submission
 Submitted (09-APR-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: LBNP
 Center clone name: CH250-269J3
 Summary Statistics
 Chemistry: Dye-primer Bodypy: inf% of reads
 Chemistry: Dye-terminator Big Dye: inf% of reads
 Consensus quality: 6461 bases at least Q40
 Consensus quality: 7875 bases at least Q30
 Consensus quality: 9630 bases at least Q20

***** NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html)
 * NOTE: The contigs are based on the application
 * of the PGI method using the Human genome (NCBI build 31)
 * as the comparative genome.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 124215: contig of 124215 bp in length.
 Location/Qualifiers
 1. 124215

FEATURES
 source


```
* 112006 114748: contig of 2743 bp in length
* 114749 114848: gap of 100 bp
* 114849 136643: contig of 21795 bp in length
* 136644 136743: gap of 100 bp
* 136744 151727: contig of 14984 bp in length
* 151728 151827: gap of 100 bp
* 151828 154062: contig of 2235 bp in length
* 154063 154162: gap of 100 bp
* 154163 157470: contig of 3308 bp in length
* 157471 157570: gap of 100 bp
* 157571 169261: contig of 11691 bp in length.
FEATURES
    source
        1..169261
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone_id="DKXY-89F23"
            /clone_id="DanioKey"
            1..11905
                /note="assembly fragment:00983
                fragment chain:1"
            12006..14620
                /note="assembly fragment:00089
                fragment chain:1"
            14721..28961
                /note="assembly fragment:00764
                fragment chain:1"
            29062..40830
                /note="assembly fragment:00578
                fragment chain:1"
            40931..76234
                /note="assembly fragment:01986
                fragment chain:1"
            76335..79519
                /note="assembly fragment:00167
                fragment chain:1"
            79620..84542
                /note="assembly fragment:00241
                fragment chain:2"
            84643..111905
                /note="assembly fragment:01564
                fragment chain:2"
            112006..114748
                /note="assembly fragment:00144
                fragment chain:2"
            114849..136643
                /note="assembly fragment:01192
                fragment chain:2"
            136744..151727
                /note="assembly fragment:00282
                fragment chain:2"
            151828..154062
                /note="assembly fragment:00104"
                fragment chain:2"
            154163..157470
                /note="assembly fragment:00200.0"
                fragment chain:2"
            157571..169261
                /note="assembly fragment:00428"
                fragment chain:2"
ORIGIN
Query Match 91.1%; Score 16.4; DB 2; Length 169261;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGACACACAGATTGCC 18
Db 5659 GTGACACACAGATTGCC 5642
RESULT 9
CR382377 214536 bp DNA linear HMG 03-APR-2004
LOCUS Danio rerio clone DKXY-238113, *** SEQUENCING IN PROGRESS ***, 28
DEFINITION unordered pieces.
```

```
ACCESSION CR382377 GI:46241759
VERSION HTG; HTGS_PHASE1.
KEYWORDS SOURCE
ORGANISM Danio rerio (zebrafish)
REFERENCE
    AUTHORS Danio rerio
    TITLE Danio rerio
    JOURNAL 1 (bases 1 to 214536)
COMMENT
    Submitted (01-Apr-2004) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    zfih-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
    On Apr 6, 2004 this sequence version replaced gi:46241599.
    ----- Genome Center
    Center: Wellcome Trust Sanger Institute
    Center code: SC
    Web site: http://www.sanger.ac.uk
    Contact: zfih-help@sanger.ac.uk
    ----- Project Information
    Center project name: zK238113
    ----- Summary Statistics
    Assembly program: XGAP4; version 4.5
    Chemistry: Dye-terminator; 100% of reads
    Consensus quality: 205798 bases at least Q40
    Consensus quality: 207592 bases at least Q30
    Consensus quality: 208725 bases at least Q20
    Insert size: 211836; sum-of-contigs
    Insert size: 170072; 7.1% error; agarose-fp
    Quality coverage: 5.74x in Q20 bases; sum-of-contigs Quality
    coverage: 8.57x in Q20 bases; agarose-fp
    -----
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 28 contigs. The true order of the pieces
    * is not known and their order in this sequence record is
    * arbitrary. Gaps between the contigs are represented as
    * runs of N, but the exact sizes of the gaps are unknown.
    * This record will be updated with the finished sequence
    * as soon as it is available and the accession number will
    * be preserved.
    *
    1 24042: contig of 24042 bp in length
    * 24043 24142: gap of 100 bp
    * 24143 45611: contig of 21469 bp in length
    * 45612 45711: gap of 100 bp
    * 45712 50249: contig of 4538 bp in length
    * 50250 50349: gap of 100 bp
    * 50350 56962: contig of 6613 bp in length
    * 56963 57062: gap of 100 bp
    * 57063 70485: contig of 13423 bp in length
    * 70486 70585: gap of 100 bp
    * 70586 89747: contig of 19162 bp in length
    * 89748 89847: gap of 100 bp
    * 89848 93674: contig of 3827 bp in length
    * 93675 93774: gap of 100 bp
    * 93775 96584: contig of 2810 bp in length
    * 96585 96684: gap of 100 bp
    * 96685 107396: contig of 10712 bp in length
    * 107397 107496: gap of 100 bp
    * 107497 145688: contig of 38172 bp in length
    * 145689 145768: gap of 100 bp
    * 145769 158179: contig of 12411 bp in length
    * 158180 158279: gap of 100 bp
    * 158280 162104: contig of 3825 bp in length
    * 162105 162205: gap of 100 bp
    * 162206 169142: contig of 6938 bp in length
    * 169143 169242: gap of 100 bp
    * 169243 171568: contig of 2326 bp in length
    * 171569 171668: gap of 100 bp
    * 171669 175613: contig of 3945 bp in length
    * 175614 175713: gap of 100 bp
    * 175714 178967: contig of 3254 bp in length
    * 178968 179067: gap of 100 bp
```

```

* 179068 181118: contig of 2051 bp in length
* 181119 181218: gap of 100 bp
* 181219 183470: contig of 2252 bp in length
* 183471 183570: gap of 100 bp
* 183571 185728: contig of 2158 bp in length
* 185729 185828: gap of 100 bp
* 185829 188116: contig of 2288 bp in length
* 188117 188216: gap of 100 bp
* 188217 190461: contig of 2245 bp in length
* 190462 190561: gap of 100 bp
* 190562 193442: contig of 2881 bp in length
* 193443 193542: gap of 100 bp
* 193543 195057: contig of 5515 bp in length
* 195058 195157: gap of 100 bp
* 195158 202015: contig of 2858 bp in length
* 202016 202115: gap of 100 bp
* 202116 204305: contig of 2190 bp in length
* 204306 204405: gap of 100 bp
* 204406 206588: contig of 2183 bp in length
* 206589 206688: gap of 100 bp
* 206689 209849: contig of 3161 bp in length
* 209850 209949: gap of 100 bp
* 209950 214536: contig of 4587 bp in length.
FEATURES
    source
    1. 214536
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /clone_lib="DKEY-238113"
        /clone_lib="DanioKey"
    1. 24042
        /note="assembly fragment:00928"
        fragment_chain:1
        24143..45611
        /note="assembly fragment:01235"
        fragment_chain:1
        45712..50249
        /note="assembly fragment:00001"
        fragment_chain:1
        50350..56962
        /note="assembly fragment:00166"
        fragment_chain:1
        57083..70485
        /note="assembly fragment:00552"
        fragment_chain:1
        70586..89747
        /note="assembly fragment:00739"
        fragment_chain:1
        89848..93674
        /note="assembly fragment:02435"
        fragment_chain:1
        93775..96584
        /note="assembly fragment:02387"
        fragment_chain:1
        96685..107396
        /note="assembly fragment:00267"
        fragment_chain:1
        107497..145668
        /note="assembly fragment:01551"
        fragment_chain:1
        145769..158179
        /note="assembly fragment:00402"
        fragment_chain:1
        158280..162104
        /note="assembly fragment:02467"
        fragment_chain:1
        162205..169142
        /note="assembly fragment:00073"
        fragment_chain:1
        169243..171568
        /note="assembly fragment:02257"
        fragment_chain:2
        171669..175613

```

```

misc_feature /note="assembly fragment:02401"
fragment_chain:2
175714..178967
/misc_feature /note="assembly fragment:02375"
fragment_chain:2
179068..181118
/misc_feature /note="assembly fragment:02343"
fragment_chain:2
181219..183470
/misc_feature /note="assembly fragment:02292"
fragment_chain:3
183571..185728
/misc_feature /note="assembly fragment:02325"
fragment_chain:3
185829..188116
/misc_feature /note="assembly fragment:02307"
fragment_chain:4
188217..190461
/misc_feature /note="assembly fragment:02313"
fragment_chain:4
190562..193442
/misc_feature /note="assembly fragment:02357"
fragment_chain:5
193543..199057
/misc_feature /note="assembly fragment:02417"
fragment_chain:5
199158..202015
/misc_feature /note="assembly fragment:02277"
202116..204305
/misc_feature /note="assembly fragment:02331"
204406..206588
/misc_feature /note="assembly fragment:02350"
206689..209849
/misc_feature /note="assembly fragment:02365"
209950..214536
/misc_feature /note="assembly fragment:02501.0"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 2; Length 214536;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GAGAACACACAGATTCCG 18

DB 209027 GTGAACACACAGATTCCG 209044

RESULT 10

AC111804 AC111804 239570 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-27013, WORKING DRAFT SEQUENCE.

```

DEFINITION AC111804
ACCESSION AC111804
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE 1 (Bases 1 to 239570)
AUTHORS Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Altschrocks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cadenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregor, G., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Lounsdar, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindrasekhar, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 239570)
Rat Genome Sequencing Consortium.
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23195895.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GODE
Center clone name: CH230-27013
Center Summary Statistics
Assembly program: Phrap; version 0.990329

Consensus quality: 227212 bases at least Q40
Consensus quality: 228362 bases at least Q30
Consensus quality: 229227 bases at least Q20
Estimated insert size: 233057; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* By the finished sequence as soon as it is available and the accession number will be preserved.
* 1 239570: contig of 239570 bp in length.

FEATURES
source

1..239570
/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-27013"

1..1150

/note="wgs end extension"

clone_end:T7"

complement(7684..8310)

/note="clone boundary"

clone_end:T7

site:

end sequence: BH338730"

complement(238848..239097)

/note="clone boundary"

clone_end:Sp6

site:

end sequence: BH338732"

ORIGIN

Query Match 91.1%; Score 16.4; DB 2; Length 239570;

Best Local Similarity 94.4%; Pred. No. 5.1e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTGCG 18

DB 111715 GAGAACACACAGATTGCG 111732

AC111225 256123 bp DNA linear HTG 13-MAY-2003

Rattus norvegicus clone CH230-109120, *** SEQUENCING IN PROGRESS

***, 3 unordered pieces.

AC111225 5 GI:30578498

HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 256123)

Muzny, D., Marle, Metzker, M., Lee, A., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyaibebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, P.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Muzny, D., Marle, Metzker, M., Lee, A., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyaibebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, P.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebragis, E., Geer, K., Gill, R., Grady, M., Gueria, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusshewa, L., Louised, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McGeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okwona, G., Olarunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanar, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 256123)
Worley, K.C.

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 256123)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24942391.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLVI
Center clone name: CH230-109120

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 237968 bases at least Q40
Consensus quality: 240525 bases at least Q30
Consensus quality: 242121 bases at least Q20
Estimated insert size: 256749; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/gcsc/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 251516: contig of 251516 bp in length
251517 251616: gap of unknown length
251617 253761: contig of 2145 bp in length
253762 253861: gap of unknown length
253862 256123: contig of 2262 bp in length.

Location/Qualifiers
1. 256123
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-109120"

1. 2094
/note="wgs_end_extension
clone_end:T7"
3849. 4622
/note="clone_boundary
clone_end:T7
site:ECORI
end sequence:BH300712"
175598. 176593
/note="wgs_contig"
246978. 247860
/note="clone_boundary
clone_end:Sp6
site:ECORI
end sequence:BH300713"
247056. 247857
/note="clone_boundary
clone_end:Sp6
site:ECORI
end sequence:BH300713"
248853. 251516
/note="wgs_end_extension
clone_end:Sp6"

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

Query Match 91.1%; Score 16.4; DB 2; Length 256123;
Best Local Similarity 94.4%; Pred. No. S.le+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCGC 18
|||||
DB 148114 GAGAACACACAGATTCAC 148037

RESULT 12
MLEPRTN8
LOCUS
DEFINITION Mycobacterium leprae strain TN complete genome; segment 8/10.
ACCESSION AL583924 AL450380
VERSION AL583924.1 GI:13093618
KEYWORDS
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

1 (bases 1 to 342300)
 Cole, S.T., Eiglmier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Ganier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, Quail, M.A., Rajandream, M.-A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Taylor, K., Whitehead, S., Woodward, J.R. and Barrell, B.G.
 Massive gene decay in the leprosy bacillus
 Nature 409 (6823), 1007-1011 (2001)
 21128732
 11234002
 2 (bases 1 to 342300)
 Parkhill, J.
 Direct Submission
 Submitted (20-Feb-2001) Submitted on behalf of the Mycobacterium leprae sequencing teams, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK Unite de Genetique Moleculaire Bacterienne, Institut Pasteur 28 rue du Docteur Roux, 75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk

Notes:
 Details of M. leprae sequencing at the Sanger Centre are available from http://www.sanger.ac.uk/projects/M_leprae/. A relational database containing the M. leprae sequences is available from <http://genolist.pasteur.fr/Leproma/>.

Location/Qualifiers
 1. 342300
 /organism="Mycobacterium leprae"
 /mol_type="genomic DNA"
 /strain="TN"
 /db_xref="taxon:1769"
 275..430
 /gene="ML1963"
 /pseudo
 /pseudo
 275..430
 /gene="ML1963"
 /note="Possible pseudogene of M. tuberculosis orthologue Rv3463 (Best blastx score 121)"
 /pseudo
 /transl_table=11
 /product="probable neuraminidase (pseudogene)"
 483..1484
 /gene="rmlB"
 /note="synonym: ML1964"
 483..1484
 /gene="rmlB"
 /note="Similar to M. tuberculosis dTDP-glucose 4,6-dehydratase rmlB Rv3464 TR:O06329 (EMBL:Z95390) (331 aa); FastA score E(): 0, 84.0% identity in 331 aa overlap, and TR:Q50556 (EMBL:U43540) (329 aa); FastA score E(): 0, 78.2% identity in 331 aa overlap, and to others e.g. Streptococcus pneumoniae dTDP-glucose-4,6-dehydratase cpn TR:O54611 (EMBL:AF030364) (349 aa); FastA score E(): 0, 59.9% identity in 334 aa overlap. Contains Pfam match to entry PF01370 Epimerase, NAD dependent epimerase/dehydratase family. Similar to ML0204, ML0751, ML1942 and ML2428"
 /codon_start=1
 /transl_table=11
 /product="dTDP-glucose 4,6-dehydratase"
 /protein_id="CAC30919.1"
 /db_xref="GI:13093619"
 /db_xref="GOA:Q9CBH7"
 /db_xref="TrEMBL:Q9CBH7"
 /translation="MTMELLVTGGAGFTGANFVNVRDHPDDPTVTLDAVTAGRHE SLADVAVRLVQGNITDTKLVRFAESDVVHFAETHADVDNALDDPEPLFNVLG TPTILEAVRHSHLHISTDEVYGDLEDEPTFTSTPNPSSPSATKAAGDMLV RAWRSYGVRAITSCNSNNYGPYQHVKEFIPTQINVLTRGRPKLYGTGVNVRDWIHV

misc_feature

DDHNSAVOILEKGQIGRTYLIGAENHNLTVLRTLLQMMGRDNDFDHVTDRVGH
 LRYAIDPTPLYNLCWAPKHFANFDEGLRATIDWYRNESWWRPLKDAVEARVEGR
 495..1424
 /gene="rmlB"
 /note="Pfam match to entry PF01370 Epimerase, NAD dependent epimerase/dehydratase family, score 470.30, E-value 1.6e-137"
 1486..2094
 /gene="rmlC"
 /note="synonym: ML1965"
 1486..2094
 /gene="rmlC"
 /note="Similar to M. tuberculosis dTDP-4-dehydrorhamnose 3,5-epimerase rmlC Rv3465 TR:O06330 (EMBL:Z95390) (202 aa); FastA score E(): 0, 75.4% identity in 199 aa overlap, and to many others e.g. Streptomyces griseus dTDP-4-dehydrorhamnose 3,5-epimerase SW:STRM_STRGR (P29783) (200 aa); FastA score E(): 0, 45.6% identity in 193 aa overlap. Contains Pfam match to entry PF00908 dTDP_sugar_isom, dTDP-4-dehydrorhamnose 3,5-epimerase."
 /codon_start=1
 /transl_table=11
 /product="dTDP-4-dehydrorhamnose 3,5-epimerase"
 /protein_id="CAC30920.1"
 /db_xref="GI:13093620"
 /db_xref="GOA:Q9X7A4"
 /db_xref="TrEMBL:Q9X7A4"
 /translation="MNVKELDISGAWETITIHSDRGMPFELWLTASRFAGFTGHRLD VRQNCSSAGVLRLGHFAQPPSQAKYTCVTGSVDFVVDIRLGSPTFGWTSIT LNSNRKTIYISLAHFLALQDNSTWYLCAGYNPAREHAICATDPLDLADWPV DGVDTLHSLKRDATAPNLDIRASGLLPTWDETQNFANLGSK"

misc_feature

1492..2025
 /gene="rmlC"
 /note="Pfam match to entry PF00908 dTDP_sugar_isom, dTDP-4-dehydrorhamnose 3,5-epimerase, score 308.80, E-value 6.5e-89"
 2139..2624
 /gene="lpqH"
 /note="synonym: ML1966"
 2139..2624
 /gene="lpqH"
 /note="Similar to M. tuberculosis 19 kDa lipoprotein antigen precursor lpqH Rv3763. Contains a possible N-terminal signal sequence and a PS00013 Prokaryotic membrane lipoprotein lipid attachment site."
 /codon_start=1
 /transl_table=11
 /product="possible lipoprotein"
 /protein_id="CAC30921.1"
 /db_xref="GI:13093621"
 /db_xref="TrEMBL:Q9X7A5"
 /translation="MRHKLIAIYAVTIMAGAAGCGGTQCAPTPSVSKTNSPTTVA SSIPDAAGSTKVTIGQPKVSGPVVCSVTGKPSIAIGDMITGVUGLEPDASVWH NAGLGLTGDVWVTAFTGVSENNATKNGNTIQTGASGVNTGGQIHKRSFEIVIC R"

gene

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

CDS

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

misc_feature

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

gene

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

CDS

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

gene

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

CDS

2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)


```

/gene="ML1968"
/Note="Possible pseudogene of M. tuberculosis paralog PE"
/pseudo
/codon_start=1
/transl_table=11
/product="PE family protein (pseudogene)"
/complement(5661..8043)
/Note="Dispersed repeat, LEPREP, copy 6"
/complement(5848..5946)
/gene="ML1969"
/pseudo
/complement(5848..5946)
/gene="ML1969"
/Note="Similar to Agrobacterium tumefaciens transposase
TR:Q44454 (EMBL:Z18270) (366 aa) fasta scores: E():
0.0002, 51.5% id in 33 aa, and to Pseudomonas putida
transposase tnpal TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta
scores: E(): 0.00088, 48.5% id in 33 aa"
/pseudo
/codon_start=1
/transl_table=11
/complement(6211..7354)
/gene="ML1970"
/pseudo
/complement(6211..7354)
/gene="ML1970"
/Note="Similar to many e.g. Cryphonectria parasitica
(Chestnut blight fungus) putative maturase TR:AAF27656
(EMBL:AF218567) (778 aa) fasta scores: E(): 7.8e-11, 29.2%
id in 216 aa"
/pseudo
/codon_start=1
/transl_table=11
/product="putative group II intron maturase-related
protein"
/complement(6674..6727)
/gene="ML1970"
/complement(7559..7698)
/gene="ML1971"
/pseudo
/complement(7559..7698)
/gene="ML1971"
/Note="Similar to Pseudomonas putida transposase tnpal
TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta scores: E():
3.4e-05, 47.8% id in 46 aa, and to Agrobacterium
tumefaciens transposase TR:Q44454 (EMBL:Z18270) (366 aa)
fasta scores: E(): 0.00026, 41.3% id in 46 aa"

Query Match      91.1%; Score 16.4; DB 1; Length 342300;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCG 18
Db 110620 GAGTACACACAGATTCG 110637

RESULT 13
AX481359
LOCUS AX481359 19 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 6 from Patent EP1225232.
ACCESSION AX481359
VERSION AX481359.1 GI:22316280
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
Rubin,B.Y. and Anderson,S.L.
AUTHORS Detection of mutations in a gene encoding lkappab
TITLE Kinase-complex-associated protein to diagnose familial dysautonomia
JOURNAL Patent: EP 1225232-A 6 24-JUL-2002;

```

```

FEATURES
Source
Rubin, Berish Y. (US) ; Anderson, Silvia L. (US)
Location/Qualifiers
1..19
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      89.9%; Score 16; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
Db 1 GAGAACACACAGATTC 16

RESULT 14
AC009112
LOCUS AC009112 64167 bp DNA linear HTG 26-JAN-2000
DEFINITION Homo sapiens chromosome 16 clone RP11-468p11, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC009112
VERSION AC009112.2 GI:6758904
KEYWORDS HTG; HTGS PHASEO.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
1 (Bases 1 to 64167)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (Bases 1 to 64167)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jan 26, 2000 this sequence version replaced gi:5685963.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

*-----*
* NOTE: This record contains 82 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 726: contig of 726 bp in length
* gap of unknown length
*
* 727 1457: contig of 731 bp in length
* gap of unknown length
*
* 1458 2142: contig of 685 bp in length
* gap of unknown length
*
* 2143 2680: contig of 538 bp in length
* gap of unknown length
*
* 2681 3386: contig of 706 bp in length
* gap of unknown length
*
* 3387 4102: contig of 716 bp in length
* gap of unknown length
*
* 4103 4654: contig of 552 bp in length
* gap of unknown length
*
* 4655 5221: contig of 567 bp in length
* gap of unknown length
*
* 5222 6170: contig of 949 bp in length
* gap of unknown length
*

```

* 6171	6805:	contig of 635 bp in length	* 31951	gap of unknown length
* 6806	7053:	gap of unknown length	* 32650:	contig of 700 bp in length
* 7054	7674:	contig of 248 bp in length	* 33308:	gap of unknown length
* 7675	8411:	gap of unknown length	* 33965:	contig of 658 bp in length
* 8412	9038:	gap of unknown length	* 34708:	gap of unknown length
* 9039	9748:	contig of 621 bp in length	* 35781:	gap of unknown length
* 9749	10370:	gap of unknown length	* 36274:	contig of 657 bp in length
* 10371	11089:	contig of 710 bp in length	* 37097:	gap of unknown length
* 11090	11781:	contig of 622 bp in length	* 37759:	contig of 823 bp in length
* 11782	12496:	gap of unknown length	* 38327:	contig of 662 bp in length
* 12497	13259:	contig of 719 bp in length	* 38328:	gap of unknown length
* 13260	13372:	contig of 692 bp in length	* 39523:	contig of 1196 bp in length
* 13373	14620:	gap of unknown length	* 40180:	gap of unknown length
* 14621	15606:	contig of 1248 bp in length	* 40799:	gap of unknown length
* 15607	16393:	gap of unknown length	* 41387:	contig of 619 bp in length
* 16394	17093:	contig of 787 bp in length	* 42390:	gap of unknown length
* 17094	17798:	gap of unknown length	* 42984:	contig of 1003 bp in length
* 17799	18473:	contig of 700 bp in length	* 43746:	gap of unknown length
* 18474	19192:	gap of unknown length	* 44488:	contig of 594 bp in length
* 19193	19908:	contig of 719 bp in length	* 45659:	gap of unknown length
* 19907	20671:	gap of unknown length	* 45790:	contig of 742 bp in length
* 20672	21341:	contig of 765 bp in length	* 46221:	gap of unknown length
* 21342	22132:	gap of unknown length	* 47164:	contig of 1171 bp in length
* 22133	22791:	contig of 791 bp in length	* 48194:	gap of unknown length
* 22792	23584:	gap of unknown length	* 49050:	contig of 1030 bp in length
* 23555	24087:	contig of 670 bp in length	* 49702:	gap of unknown length
* 24088	24798:	gap of unknown length	* 49883:	contig of 856 bp in length
* 24799	25531:	contig of 533 bp in length	* 51039:	gap of unknown length
* 25532	25668:	gap of unknown length	* 52165:	contig of 1156 bp in length
* 25669	26491:	contig of 137 bp in length	* 53296:	gap of unknown length
* 26492	27242:	gap of unknown length	* 54764:	contig of 1131 bp in length
* 27243	27842:	contig of 823 bp in length	* 55985:	gap of unknown length
* 27843	28475:	gap of unknown length	* 57280:	contig of 1221 bp in length
* 28476	29146:	contig of 751 bp in length	* 58306:	gap of unknown length
* 29147	29734:	gap of unknown length	* 59481:	contig of 1026 bp in length
* 29735	31298:	contig of 600 bp in length	* 60351:	gap of unknown length
* 31299	31950:	gap of unknown length	* 63313:	contig of 870 bp in length
			* 64167:	gap of unknown length
			* 63314	contig of 2962 bp in length
				Location/Qualifiers
				contig of 854 bp in length.

Query Match 88.9%; Score 16; DB 2; Length 64167;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 |||||
 Db 21230 GAGAACACACAGATTC 21245

RESULT 15
 LOCUS AX676048 66479 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 1 from Patent WO02059381.
 ACCESSION AX676048
 VERSION AX676048.1 GI:29333739

KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Slangenaupt,S. and Gusella,J.F.
 AUTHORS Gene for identifying individuals with familial dysautonomia
 TITLE Patent: WO 02059381-A 1 01-AUG-2002;
 JOURNAL The General Hospital Corporation (US)

FEATURES
 Location/Qualifiers
 source
 1..66479
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 88.9%; Score 16; DB 6; Length 66479;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 |||||
 Db 33979 GAGAACACACAGATTC 33994

RESULT 16
 LOCUS AL359692/c 78376 bp DNA linear PRI 18-JUL-2001
 DEFINITION Human DNA sequence from clone RP11-3J11 on chromosome 9, complete
 sequence.

ACCESSION AL359692
 VERSION AL359692.9 GI:14970800
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 78376)
 Direct Submission
 Submitted (18-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clones@sanger.ac.uk

COMMENT
 On Jul 19, 2001 this sequence version replaced gi:14715342.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:,
 SWISSPROT; Tr: TREMBL; Wp:, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-3J11 is from the library RPCI-11.1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-3J11 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-3J11 is at 1 in this sequence. The
 true left end of clone RP11-115J22 is at 76377 in this sequence.
 The true right end of clone RP11-339N8 is at 76382 in this
 sequence.

FEATURES

Location/Qualifiers

source
 1..78376
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-3J11"
 /clone_lib="RPCI-11.1"
 repeat_region
 7..829
 /note="LIPB1 repeat: matches 5343..6155 of consensus"
 991..1328
 /note="MER46B repeat: matches 3..231 of consensus"
 2199..2234
 /note="18 copies 2 mer tt 86% conserved"
 3191..3349
 /note="AluSg/x repeat: matches 153..311 of consensus"
 4250..4537
 /note="AluSx repeat: matches 1..288 of consensus"
 5314..5579
 /note="AluJb repeat: matches 1..299 of consensus"
 5693..5863
 /note="L1MC4 repeat: matches 6481..6679 of consensus"
 5881..6013
 /note="AluJb repeat: matches 1..135 of consensus"
 6533..6644
 /note="MIR repeat: matches 92..211 of consensus"
 7645..7979
 /note="MLT1J repeat: matches 3..363 of consensus"
 9165..9774
 /note="MER21B repeat: matches 179..787 of consensus"
 9775..9848
 /note="MER84 repeat: matches 1..76 of consensus"
 9849..10232
 /note="MER84 repeat: matches 3..373 of consensus"
 10233..10535
 /note="AluSx repeat: matches 1..302 of consensus"
 10536..10668
 /note="MER84 repeat: matches 373..508 of consensus"
 10669..10846
 /note="MER21B repeat: matches 7..185 of consensus"
 11168..11901
 /note="LIPAL4 repeat: matches 5403..6149 of consensus"
 12172..12468
 /note="AluSg repeat: matches 1..297 of consensus"
 13074..13362
 /note="AluSx repeat: matches 3..294 of consensus"
 13411..13533
 /note="MIR repeat: matches 117..239 of consensus"
 13616..13669
 /note="L2 repeat: matches 2445..2498 of consensus"
 13733..13841
 /note="MER53 repeat: matches 89..188 of consensus"

```

repeat_region 13996. .14113
/note="AluSg/x repeat: matches 178. .295 of consensus"
repeat_region 14769. .15142
/note="WSTA repeat: matches 1. .423 of consensus"
repeat_region 15274. .15333
/note="WERA repeat: matches 14. .74 of consensus"
repeat_region 15480. .15600
/note="MER5B repeat: matches 54. .175 of consensus"
repeat_region 15640. .15689
/note="WSTA repeat: matches 1. .49 of consensus"
repeat_region 15861. .16180
/note="AluB repeat: matches 1. .302 of consensus"
repeat_region 17189. .17296
/note="MIR repeat: matches 100. .206 of consensus"
repeat_region 17614. .17738
/note="MER5B repeat: matches 54. .174 of consensus"
repeat_region 17769. .18234
/note="L1MD1 repeat: matches 5748. .6223 of consensus"
repeat_region 18235. .18529
/note="AluJ repeat: matches 1. .292 of consensus"
repeat_region 18530. .19071
/note="L1MD1 repeat: matches 5238. .5748 of consensus"
repeat_region 19132. .19432
/note="AluSp repeat: matches 1. .313 of consensus"
repeat_region 19441. .19472
/note="16 copies 2 mer tt 84% conserved"
repeat_region 19498. .19859
/note="THE1A repeat: matches 1. .353 of consensus"
repeat_region 19860. .21021
/note="THE1A-internal repeat: matches 427. .1580 of consensus"
repeat_region 21022. .21314
/note="AluSg repeat: matches 1. .302 of consensus"
repeat_region 21315. .21711
/note="THE1A-internal repeat: matches 29. .427 of consensus"
repeat_region 21718. .21818
/note="THE1B repeat: matches 265. .364 of consensus"
repeat_region 21813. .21845
/note="Alu repeat: matches 3. .35 of consensus"
repeat_region 21878. .22184
/note="AluSg repeat: matches 1. .311 of consensus"
repeat_region 22578. .22715
/note="MIR repeat: matches 74. .210 of consensus"
repeat_region 22693. .22848
/note="L2 repeat: matches 2572. .2745 of consensus"
repeat_region 24368. .24434
/note="MIR repeat: matches 71. .137 of consensus"
repeat_region 25259. .25338
/note="L2 repeat: matches 2605. .2695 of consensus"
repeat_region 25474. .25775
/note="AluSx repeat: matches 1. .301 of consensus"
misc_feature 26871. .27694
/note="CpG island"
/evidence=not_experimental
repeat_region 28479. .28806
/note="L2 repeat: matches 2157. .2496 of consensus"
repeat_region 28807. .29114
/note="AluS1 repeat: matches 1. .305 of consensus"
repeat_region 29115. .29256
/note="L2 repeat: matches 1963. .2157 of consensus"
repeat_region 29748. .30050
/note="AluSg repeat: matches 1. .303 of consensus"
repeat_region 30475. .30548
/note="MIR repeat: matches 33. .107 of consensus"
repeat_region 30654. .30701
/note="24 copies 2 mer ac 95% conserved"
repeat_region 31870. .32343
/note="MLTID repeat: matches 16. .505 of consensus"
repeat_region 33199. .33414
/note="L2 repeat: matches 2116. .2353 of consensus"
repeat_region 33415. .33568
/note="MER5B repeat: matches 186. .340 of consensus"

```

```

repeat_region 33658. .33918
/note="L2 repeat: matches 2403. .2710 of consensus"
repeat_region 37241. .37362
/note="Charlie4 repeat: matches 1823. .1956 of consensus"
repeat_region 37669. .37895
/note="MER20 repeat: matches 1. .218 of consensus"
repeat_region 38589. .38875
/note="AluSg repeat: matches 1. .310 of consensus"
repeat_region 39767. .39947
/note="WSTB repeat: matches 1. .183 of consensus"
misc_feature 40006. .40255
/note="match: STS: Em:G14836"
misc_feature 40477. .40798
/note="match: STS: Em:G21606"
repeat_region 41261. .41395
/note="L1M2 repeat: matches 6041. .6161 of consensus"
repeat_region 41739. .41805
/note="TIGER2 repeat: matches 2653. .2718 of consensus"
repeat_region 41806. .42093
/note="AluB repeat: matches 1. .300 of consensus"
repeat_region 42094. .43429
/note="TIGER2 repeat: matches 1302. .2653 of consensus"
repeat_region 43429. .44232
/note="TIGER2 repeat: matches 297. .1113 of consensus"
repeat_region 44353. .44664

Query Match 88.9%; Score 16; DB 9; Length 78376;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||
Db 71868 GAGAACACACAGATTC 71853

RESULT 17
AC135161/c AC135161 88701 bp DNA linear PLN 15-MAR-2004
LOCUS Medicago truncatula clone mth2-30k24, complete sequence.
DEFINITION
AC135161
ACCESSION
VERSION AC135161.11 GI:45434547
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-30k24
Unpublished
REFERENCE 2 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (08-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (06-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 4 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (21-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

OK 73019, USA
5 (bases 1 to 88701)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (12-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 88701)
Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (15-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Mar 15, 2004 this sequence version replaced gi:45384562.

Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES
source
Location/Qualifiers
1..88701
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-30k24"
/clone_lib="Medicago truncatula BAC library H2"

ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 88701;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTCG 17
Db 23358 AGAACACACAGATTTCG 23343

RESULT 18
CR382130_35/c
WPCOMMENT
Sequence split into 37 fragments LOCUS CR382130 Accession CR382130

Fragment Name	Begin	End
CR382130_00	1	110000
CR382130_01	100001	210000
CR382130_02	200001	310000
CR382130_03	300001	410000
CR382130_04	400001	510000
CR382130_05	500001	610000
CR382130_06	600001	710000
CR382130_07	700001	810000
CR382130_08	800001	910000
CR382130_09	900001	1010000
CR382130_10	1000001	1100000
CR382130_11	1100001	1210000
CR382130_12	1200001	1310000
CR382130_13	1300001	1410000
CR382130_14	1400001	1510000
CR382130_15	1500001	1610000
CR382130_16	1600001	1710000
CR382130_17	1700001	1810000
CR382130_18	1800001	1910000
CR382130_19	1900001	2010000
CR382130_20	2000001	2110000
CR382130_21	2100001	2210000
CR382130_22	2200001	2310000
CR382130_23	2300001	2410000
CR382130_24	2400001	2510000
CR382130_25	2500001	2610000
CR382130_26	2600001	2710000
CR382130_27	2700001	2810000
CR382130_28	2800001	2910000

CR382130_29 2900001 3010000
CR382130_30 3000001 3110000
CR382130_31 3100001 3210000
CR382130_32 3200001 3310000
CR382130_33 3300001 3410000
CR382130_34 3400001 3510000
CR382130_35 3500001 3610000
CR382130_36 3600001 3632272
Continuation (36 of 37) of CR382130 from base 3500001 (CR382130 Yarrowia lipolytica chron

Query Match 88.9%; Score 16; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTC 16
Db 77927 GAGACACACAGATTTC 77912

RESULT 19
CR382133_06/c
WPCOMMENT
Sequence split into 13 fragments LOCUS CR382133 Accession CR382133

Fragment Name	Begin	End
CR382133_00	1	110000
CR382133_01	100001	210000
CR382133_02	200001	310000
CR382133_03	300001	410000
CR382133_04	400001	510000
CR382133_05	500001	610000
CR382133_06	600001	710000
CR382133_07	700001	810000
CR382133_08	800001	910000
CR382133_09	900001	1010000
CR382133_10	1000001	1110000
CR382133_11	1100001	1210000
CR382133_12	1200001	1249565

Continuation (7 of 13) of CR382133 from base 600001 (CR382133 Debaryomyces hansenii chron

Query Match 88.9%; Score 16; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTCG 17
Db 17211 AGAACACACAGATTTCG 17196

RESULT 20
AC011700/c
LOCUS
DEFINITION Homo sapiens Chromosome 1 BAC RP11-478J18 (Roswell Park Cancer
Institute Human BAC Library) complete sequence.

AC011700 154732 bp DNA linear PRI 13-JUN-2000
AC011700
VERSION AC011700.5 GI:8439783
KEYWORDS HTG.
SOURCE Homo sapiens (human)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 154732)
Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
Bodot,B., Buck,J., Bowler,S., Brooks,A., Bunay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hoques,M.,
Holloway,C., Hosak,H., Issar,A., Jackson,L.E., Jackson,L., Jia,Y.,
Jones,M., Kelly,S., Kneitz,S., Kondejewski,N., Kong,Y., Kovar,C.,
Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,
Logan,O., Lu,J., Lucier,R., Marondel,I., Martin,R., Martinez,C.,
McLeod,M.P., Mei,G., Merscher,S., Miller,A., Montgomery,K.T.,


```

repeat_region      26242..26300
                    /rpt_family="A-rich"
repeat_region      26918..26970
                    /rpt_family="(TA)n"
repeat_region      31000..31061
                    /rpt_family="(CA)n"
repeat_region      31068..31090
                    /rpt_family="(TC)n"
repeat_region      31548..31587
                    /rpt_family="(T)n"
misc_feature       31825..32144
                    /note="Region: H.sapiens CpG island DNA genomic MseI
                    fragment, clone 33a11, Z60720"
misc_feature       32403..32608
                    /note="Region: H.sapiens CpG island DNA genomic MseI
                    fragment, clone 33a11, Z60719"
repeat_region      32655..33310
                    /rpt_family="(TC)n"
repeat_region      33447..33626
                    /rpt_family="GA-rich"
misc_feature       33899..34029
                    /function="Low coverage"
repeat_region      33973..34018
                    /rpt_family="(CCCCCA)n"
repeat_region      34145..34172
                    /rpt_family="(T)n"
repeat_region      36959..37011
                    /rpt_family="(CCA)n"
misc_feature       37522..37889
                    /note="Region: ox03f03.s1 Homo sapiens cDNA, AI033409"
repeat_region      complement(39640..39663)

Query Match      88.9%; Score 16; DB 9; Length 154732;
Best Local Similarity 100.0%; Pred. No. 8.6e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTC 16
Db 87079 GAGACACACAGATTC 87064
|||||
|||||

RESULT 21
AC129818/c
LOCUS              157381 bp      DNA      linear      HTG 20-NOV-2002
DEFINITION         Rattus norvegicus clone CH230-392M8, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION          AC129818
VERSION            GI:25137966
KEYWORDS           HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE             Rattus norvegicus
ORGANISM           Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE           1 (bases 1 to 157381)
AUTHORS            Muzny,D,Marle,E., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
                    Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
                    Ayalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,
                    Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
                    Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
                    Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
                    Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
                    Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
                    Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
                    Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
                    Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
                    Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
                    Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
                    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
                    Fraser,C.M., Gabis,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
                    Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
                    Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
                    Harvey,Y., Havlax,P., Hawes,A., Henderson,N., Hernandez,J.,
                    Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
                    Hollins,B., Howells,S., Hulyx,S., Hume,J., Iclebird,D., Jackson,A.,
                    Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
                    Karpathy,S., Kelly,S., Khan,L., Khan,Z., King,L., Kovar,C.,
                    Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
                    Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
                    Lorenshewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
                    Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
                    Margum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
                    Mathewiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
                    Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
                    Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
                    Narkervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
                    Nwankwelenh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
                    Pasternak,S., Paul,H., Perez,A., Popovic,D., Primus,E., Pu,L., L.,
                    Plopper,F., Pindexter,A., Popovich,D., Reeves,K., Regier,M.A., Reigh,R.,
                    Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Richards,S., Riggs,F.,
                    Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Ruiz,S.J.,
                    Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
                    Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
                    Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
                    Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
                    Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
                    Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
                    Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
                    Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
                    Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
                    Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
                    Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
                    Niederhausern,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
                    Weinstock,G. and Gibbs,R.A.
                    Direct Submission
                    2 (bases 1 to 157381)
                    Worely,K.C.
                    Direct Submission
                    Submitted (03-AUG-2002) Human Genome Sequencing Center, Department
                    of Molecular and Human Genetics, Baylor College of Medicine, One
                    Baylor Plaza, Houston, TX 77030, USA
                    3 (bases 1 to 157381)
                    Rat Genome Sequencing Consortium.
                    Direct Submission
                    Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
                    of Molecular and Human Genetics, Baylor College of Medicine, One
                    Baylor Plaza, Houston, TX 77030, USA
                    On Nov 20, 2002 this sequence version replaced gi:23907949.
                    The sequence in this assembly is a combination of BAC based reads
                    in the feature table below represents a scaffold in the Atlas
                    assembly (a 'contig-scaffold'). Within each contig described
                    individual sequence contigs are ordered and oriented, and separated
                    by sized gaps filled with Ns to the estimated size. The sequence
                    may extend beyond the ends of the clone and there may be sequence
                    contigs within a contig-scaffold that consist entirely of whole
                    genome shotgun sequence reads. Both end sequences and whole genome
                    shotgun sequence only contigs will be indicated in the feature
                    table.
                    ----- Genome Center
                    Center: Baylor College of Medicine
                    Center code: BCM
                    Web site: http://www.hgsc.bcm.tmc.edu/
                    Contact: hgsc-help@bcm.tmc.edu
                    ----- Project Information
                    Center project name: GWTE
                    Center clone name: CH230-392M8
                    ----- Summary Statistics
                    Assembly program: Phrap; version 0.990329
                    Consensus quality: 153526 bases at least Q40
                    Consensus quality: 154705 bases at least Q30
                    Consensus quality: 155316 bases at least Q20
                    Estimated insert size: 156048; sum-of-contigs estimation
                    Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

```

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 55243: contig of 55243 bp in length
 55244 55343: gap of unknown length
 55344 153147: contig of 97804 bp in length
 153148 153247: gap of unknown length
 153248 157381: contig of 4134 bp in length.

FEATURES

source
 1. 157381
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-392M8"
 misc_feature
 1. 1061
 /note="wgs end_extension"
 clone_end:Sp6"
 1117..1750
 /note="clone_boundary"
 clone_end:Sp6"
 misc_feature
 5207..5765
 /note="clone_boundary"
 clone_end:T7
 site:
 end_sequence:BZ276820"
 5207..5765
 /note="clone_boundary"
 clone_end:T7
 site:
 end_sequence:BZ276819"

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 157381;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 Db 4779 GAGAACACACAGATTC 4764
 |||||

RESULT 22
 AC007341 168612 bp DNA linear PRI 27-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-381L15, complete sequence.
 DEFINITION
 AC007341
 VERSION AC007341.10 GI:29294011
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 2 (bases 1 to 168612)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 3 (bases 1 to 168612)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 4 (bases 1 to 168612)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

AUTHORS

TITLE
 JOURNAL

REFERENCE
 AUTHORS

TITLE
 JOURNAL

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

FEATURES
 source

1. 168612
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-381L15"

ORIGIN

Query Match 88.9%; Score 16; DB 9; Length 168612;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 Db 61304 GAGAACACACAGATTC 61319
 |||||

RESULT 23
 AC007333 180837 bp DNA linear PRI 27-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-142G1, complete sequence.
 DEFINITION
 AC007333
 VERSION AC007333.8 GI:29294008
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 2 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 3 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 4 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 5 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

DOE Joint Genome Institute.

Direct Submission
 Submitted (04-DEC-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

4 (bases 1 to 168612)
 DOE Joint Genome Institute.

Direct Submission
 Submitted (05-DEC-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

5 (bases 1 to 168612)
 DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.

Direct Submission
 Submitted (27-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA

On Mar 27, 2003 this sequence version replaced GI:17352398.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center and Los Alamos
 National Laboratory
 www-shgc.stanford.edu

Quality: Phrap Quality >=40 100% of Sequence;
 Estimated Total Number of Errors is 0.7.

Location/Qualifiers
 1. 168612
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-381L15"

ORIGIN

Query Match 88.9%; Score 16; DB 9; Length 168612;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 Db 61304 GAGAACACACAGATTC 61319
 |||||

RESULT 23
 AC007333 180837 bp DNA linear PRI 27-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-142G1, complete sequence.
 DEFINITION
 AC007333
 VERSION AC007333.8 GI:29294008
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 2 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 3 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 4 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE 5 (bases 1 to 180837)
 AUTHORS Bruce, D., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Meink, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

AUTHORS	Mundt, M.O.	Direct Submission	Submitted (05-OCT-2001) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, USA	REFERENCE	5 (bases 1 to 180837)	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.	
AUTHORS	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.	Direct Submission	Submitted (27-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	TITLE	On Mar 27, 2003 this sequence version replaced gi:15284264.	COMMENT	Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory www.shgc.stanford.edu Quality: Phrap Quality >=40 100% of Sequence; Estimated Total Number of Errors is 0. NOTE: BACTERIAL TRANSPOSON excised at 77936.
FEATURES	source	Location/Qualifiers	1..180837 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosomes="16" /clone="RP11-142G1"	ORIGIN	Query Match 88.9%; Score 16; DB 9; Length 180837; Best Local Similarity 100.0%; Pred. No. 8.6e+02; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	GAGAACACAGATTC 16		Db	31939	GAGAACACAGATTC 31924	
RESULT 24	AC111389	LOCUS	AC111389 214026 bp DNA linear HTG 13-MAY-2003	DEFINITION	Rattus norvegicus clone CH230-143J16, WORKING DRAFT SEQUENCE, 5 unordered pieces.	AC111389	
VERSION	AC111389	GI:30578959	HTG: HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.	KEYWORDS	Rattus norvegicus (Norway rat)	ORGANISM	
REFERENCE	AUTHORS	1 (bases 1 to 214026)	Muzny, D. Marie, Merzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,				

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 60867: contig of 60867 bp in length
 * 60868 60967: gap of unknown length
 * 60968 20854: contig of 14887 bp in length
 * 20955 21126: gap of unknown length
 * 20955 21126: contig of 1172 bp in length
 * 21127 21226: gap of unknown length
 * 21127 21275: contig of 1531 bp in length
 * 21275 21275: gap of unknown length
 * 21275 214026: contig of 1169 bp in length.

FEATURES

source

Location/Qualifiers
 1..214026
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-143J16"
 1..1534
 /note="wgs_contig"

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 214026;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAGAACACAGATTC 16
 |||||
 Db 99497 GAGAACACAGATTC 99512

RESULT 25
 AC094494 Rattus norvegicus clone CH230-4D21, *** SEQUENCING IN PROGRESS ***
 LOCUS 2 unordered pieces.
 DEFINITION AC094494.7 GI:30466802
 VERSION HTG: HTGS PHASE1: HTGS DRAFT: HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus

REFERENCE 1 (bases 1 to 236172)
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Deigado, O., Denson, S., Derano, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganter, R., Garcia, A., Garner, T., Garza, M., Gabrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerrero, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, N., Hoggues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louised, H., Lorado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, N., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekelemeh, O., Okunolu, G., Olariunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfrankoch, C., Popper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snajds, D., Sneed, A., Sodergren, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission

Unpublished
 2 (bases 1 to 236172)

Worley, K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236172)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24942238.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GATN

Center clone name: CH230-4D21

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 220342 bases at least Q40

Consensus quality: 223840 bases at least Q30

Consensus quality: 225944 bases at least Q20

Estimated insert size: 235672; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 234509: contig of 234509 bp in length
 * 234510 234609: gap of unknown length
 * 234610 236172: contig of 1563 bp in length.

FEATURES

source

1. .236172
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-4D21"
 1. .1828
 /note="wgs_contig"

ORIGIN

Query Match 88.9%; Score 16; DB 2: Length 236172;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTC 16

Db 54358 GAGAACACAGATTC 54373

RESULT 26

AC126106/c

LOCUS

DEFINITION AC126106 251032 bp DNA linear HTG 10-MAY-2003
 Rattus norvegicus clone CH230-2604, *** SEQUENCING IN PROGRESS ***,
 2 unordered pieces.

ACCESSION

AC126106 GI:30522483

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 251032)

Muzny,D,Marie, Metzker,M, Lee, A., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,I., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenschwartz,L., Loulsgaard,H., Lozano,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mathew,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Popper, F., Poudexter, A., Popovic, D., Primus, E., Pl, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,
 Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 251032)

Worley, K.C.

Direct Submission

Submitted (03-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 251032)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23264220.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GQCN

Center clone name: CH230-2604

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 243807 bases at least Q40

Consensus quality: 245909 bases at least Q30

Consensus quality: 247143 bases at least Q20

Estimated insert size: 258832; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one 'clone'.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 249354: contig of 249354 bp in length

* 249355 249454: gap of unknown length
 * 249455 251032: contig of 1578 bp in length.

FEATURES

source
 1. 251032
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-2604"
 743. .1295
 /note="clone_boundary
 clone_end:T7
 site:ECORI
 end_sequence:BH334758"
 complement(245981..246654)
 /note="clone_boundary
 clone_end:Sp6
 site:ECORI
 end_sequence:BH334759"

misc_feature

misc_feature

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 251032;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAACACCAAGATTC 16

Db 163240 GAGACACCAAGATTC 163225

RESULT 27

AC094225

LOCUS Rattus norvegicus clone CH230-3B10, WORKING DRAFT SEQUENCE, 3
 DEFINITION unordered pieces.

ACCESSION

AC094225 GI:30467636

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 292516)
 Muzny,D.M,Marie, Metzker,M, Lee, Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Derano,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,T., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gregeorge,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Huliyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Louleghed,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwaokemele,O., Okwunu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
 Puaro,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snajds,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 292516)

Worley,K.C.

Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 292516)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24819436.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center of Medicine

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GAPS

Center clone name: CH230-3B10

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 249308 bases at least Q40

Consensus quality: 250763 bases at least Q30

Consensus quality: 251921 bases at least Q20

Estimated insert size: 263955; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

```

* 1 287193: contig of 287193 bp in length
* 287194 287293: gap of unknown length
* 287294 288662: contig of 1369 bp in length
* 288663 288762: gap of unknown length
* 288763 292516: contig of 3754 bp in length.
FEATURES
    Location/Qualifiers
        source
            1..292516
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-3B10"
        misc_feature
            1..2212
            /note="wgs contig"
        misc_feature
            45578..48328
            /note="wgs contig"
        misc_feature
            complement(285431..286059)
            /note="clone boundary"
            clone_end:r7
            site:EcORI
            end_sequence:BH307153"

ORIGIN
Query Match      88.9%; Score 16; DB 2; Length 292516;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAAAGATTC 16
    |||||
Db 135976 GAGACACACAAAGATTC 135991

RESULT 28
AX340908
LOCUS AX340908 485 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 1155 from Patent WO0196388.
ACCESSION AX340908
VERSION AX340908.1 GI:18136890
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang, Y., Harlocker, S.I. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0196388-A 1155 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
    source
        1..485
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match      85.6%; Score 15.4; DB 6; Length 485;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAAAGATTCG 18
    |||||
Db 48 AGCACACACAAAGATTCG 64

RESULT 29
AR396508
LOCUS AR396508 534 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2523 from patent US 6617156.
ACCESSION AR396508
VERSION AR396508.1 GI:40126069
KEYWORDS Unknown.
ORGANISM Unknown.

* 1 287193: contig of 287193 bp in length
* 287293: gap of unknown length
* 287294 288662: contig of 1369 bp in length
* 288663 288762: gap of unknown length
* 288763 292516: contig of 3754 bp in length.
FEATURES
    Location/Qualifiers
        source
            1..292516
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-3B10"
        misc_feature
            1..2212
            /note="wgs contig"
        misc_feature
            45578..48328
            /note="wgs contig"
        misc_feature
            complement(285431..286059)
            /note="clone boundary"
            clone_end:r7
            site:EcORI
            end_sequence:BH307153"

ORIGIN
Query Match      85.6%; Score 15.4; DB 6; Length 534;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAAAGATTCG 18
    |||||
Db 465 AGAACACACAAAGATTCG 481

RESULT 30
AY344560/c
LOCUS AY344560 886 bp DNA linear VRT 01-AUG-2004
DEFINITION Sinogastromyzon puliensis haplotype tw4 control region, complete
sequence; mitochondrial.
ACCESSION AY344560
VERSION AY344560.1 GI:38098435
KEYWORDS mitochondrion Sinogastromyzon puliensis
ORGANISM Sinogastromyzon puliensis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.
REFERENCE 1 (bases 1 to 886)
AUTHORS Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
TITLE The Effects of Glaciation on the Dispersion Pattern of the
Endangered Species, Sinogastromyzon puliensis (Cypriniformes:
Balitoridae), in Southwest Taiwan
JOURNAL Unpublished
AUTHORS Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu
Road, Hsinchu, Taiwan 300, Republic of China
FEATURES
    source
        1..886
        /organism="Sinogastromyzon puliensis"
        /organelle="mitochondrion"
        /mol_type="genomic DNA"
        /db_xref="taxon:241472"
        /haplotype="tw4"
    misc_feature
        1..886
        /note="control region"

ORIGIN
Query Match      85.6%; Score 15.4; DB 5; Length 886;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAAAGATTCG 17
    |||||
Db 177 GAGAACACACATATTCG 161

RESULT 31
AY344557/c
LOCUS AY344557 887 bp DNA linear VRT 01-AUG-2004
DEFINITION Sinogastromyzon puliensis haplotype tw1 control region, complete
sequence; mitochondrial.
ACCESSION AY344557
VERSION AY344557.1 GI:38098432
KEYWORDS

```

Matches	16;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	GAGAACAAACAAGATTTCG	17						
Db	178	GAGAACAAACAATTTCCG	162						
RESULT 33									
AY344561/c									
LOCUS	AY344561	887 bp	DNA	linear	VRT 01-AUG-2004				
DEFINITION	Sinogastromyzon pulliensis haplotype tw5 control region, complete sequence; mitochondrial.								
ACCESSION	AY344561								
VERSION	AY344561.1	GI:38098436							
KEYWORDS									
SOURCE									
ORGANISM									
	mitochondrion Sinogastromyzon pulliensis								
	Sinogastromyzon pulliensis								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;								
	Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.								
REFERENCE	1 (bases 1 to 887)								
AUTHORS	Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.								
TITLE	The Effects of Glaciation on the Dispersion Pattern of the Endangered Species, Sinogastromyzon pulliensis (Cypriniformes: Balitoridae), in Southwest Taiwan								
JOURNAL	Unpublished								
AUTHORS	2 (bases 1 to 887)								
TITLE	Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.								
JOURNAL	Direct Submission								
FEATURES	Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu Road, Hsinchu, Taiwan 300, Republic of China								
source	Location/Qualifiers								
	1..887								
	/organism="Sinogastromyzon pulliensis"								
	/mol_type="genomic DNA"								
	/db_xref="taxon:241472"								
	/haplotype="tw5"								
misc_feature	1..887								
	/note="control region"								
ORIGIN									
	Query Match	85.6%;	Score 15.4;	DB 5;	Length 887;				
	Best Local Similarity	94.1%;	Pred. No. 2.2e+03;						
Matches	16;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	GAGAACAAACAAGATTTCG	17						
Db	178	GAGAACAAACAATTTCCG	162						
RESULT 34									
AY344562/c									
LOCUS	AY344562	887 bp	DNA	linear	VRT 01-AUG-2004				
DEFINITION	Sinogastromyzon pulliensis haplotype tw6 control region, complete sequence; mitochondrial.								
ACCESSION	AY344562								
VERSION	AY344562.1	GI:38098437							
KEYWORDS									
SOURCE									
ORGANISM									
	mitochondrion Sinogastromyzon pulliensis								
	Sinogastromyzon pulliensis								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;								
	Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.								
REFERENCE	1 (bases 1 to 887)								
AUTHORS	Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.								
TITLE	The Effects of Glaciation on the Dispersion Pattern of the Endangered Species, Sinogastromyzon pulliensis (Cypriniformes: Balitoridae), in Southwest Taiwan								
JOURNAL	Unpublished								
AUTHORS	2 (bases 1 to 887)								
TITLE	Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.								
JOURNAL	Direct Submission								

JOURNAL Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu Road, Hsinchu, Taiwan 300, Republic of China

FEATURES
source Location/Qualifiers
1..887
/organism="Sinogastromyzon puliensis"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:241472"
/haplotype="tw6"
misc_feature
1..887
/note="control region"

ORIGIN
Query Match 85.6%; Score 15.4; DB 5; Length 887;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTCG 17
|||||
Db 178 GAGACACACATATTTCG 162

RESULT 35
AY344558/c
LOCUS
DEFINITION Sinogastromyzon puliensis haplotype tw2 control region, complete sequence; mitochondrial.
ACCESSION AY344558
VERSION AY344558.1 GI:38098433
KEYWORDS
SOURCE Sinogastromyzon puliensis
ORGANISM Sinogastromyzon puliensis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.
REFERENCE
1 (bases 1 to 888)
Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
The Effects of Glaciation on the Dispersion Pattern of the Endangered Species, Sinogastromyzon puliensis (Cypriniformes: Balitoridae), in Southwest Taiwan
Unpublished
2 (bases 1 to 888)
Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
Direct Submission
Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu Road, Hsinchu, Taiwan 300, Republic of China

FEATURES
source Location/Qualifiers
1..888
/organism="Sinogastromyzon puliensis"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:241472"
/haplotype="tw2"
misc_feature
1..888
/note="control region"

ORIGIN
Query Match 85.6%; Score 15.4; DB 5; Length 888;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTCG 17
|||||
Db 178 GAGACACACATATTTCG 162

RESULT 36
J8RTMPB
LOCUS
DEFINITION Jaagsiekte sheep retrovirus pol protein (pol) Gene, partial cds.
ACCESSION M65014
VERSION M65014.1 GI:331345
KEYWORDS pol polyprotein.

SOURCE
ORGANISM Ovine pulmonary adenocarcinoma virus
REFERENCE Ovine pulmonary adenocarcinoma virus
1 (bases 1 to 1524)
AUTHORS Viruses; Retrovird viruses; Retroviridae; Betaretrovirus.
TITLE York, D.F., Vigne, R., Verwoerd, D.W. and Querat, G.
Isolation, identification, and partial cDNA cloning of genomic RNA of jaagsiekte retrovirus, the etiological agent of sheep pulmonary adenomatosis
J. Virol. 65 (9), 5061-5067 (1991)
MEDLINE 91333055
PUBMED 1651422
COMMENT Original source text: Jaagsiekte sheep retrovirus cDNA to genomic RNA.

FEATURES
source Location/Qualifiers
1..1524
/organism="Ovine pulmonary adenocarcinoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11746"
/cell_line="fluid of sheep carcinoma"
1..1524
/gene="pol"
<1..>1524
/gene="pol"
/note="gene products RNase H and integrase"
/codon_start=3
/product="pol protein"
/protein_id="AAA46253.1"
/db_xref="GI:555225"
/translation="QTHFVNYLGSFLYPRVYNTQVLKLOTDLKTLNDPQKLGLDINWIRPYLKLPTLQPLFDILKQSDPASPELTLSEGRATLQSTEEAIRQQOITYCDYQFGMEPPFCVYALQEQDWLFQFSDNWSIAFNPFGQITHHYPSDKLLQSPASHAFIPKVRQPIPEATLIFTDSSNGTAAIINHQTYAGTSFSSQVVELFAVHQALLTVPTFNLTDSYVVGALMIETPIIGTTSPEVLNFTLIQQVLHCRPCFPFGHTRAHSTLPGALVOGNTADVLTQVVFQSAIDAAKSHDLHQNSHLRLQFKISREAAQIVKSCSTCPOFFVLQPYGVNRPGLRNPHLQTDVTHIPQFGRLLKYVHVSIDTFNSFLMASLTGSTRHCITQHLFCFSTSGIPOTLKTNDGPGYTSRFSQFCLSFQIHHKTGI"

ORIGIN
Query Match 85.6%; Score 15.4; DB 14; Length 1524;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTCG 18
|||||
Db 524 AGAACACACAGATTTCG 540

RESULT 37
AK109670/c
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:002-145-B01, full insert sequence.
ACCESSION AK109670
VERSION AK109670.1 GI:32994879
KEYWORDS FLI cDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

REFERENCE
1
The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team., Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,

Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M., and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PubMed 12669764

2 (bases 1 to 1664)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imoto,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Nami,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Oneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yanada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602 Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Nami,T.,
Oneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kouda,M.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1. .1664
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

/clone="002-145-B01"
ORIGIN
Query Match 85.6%; Score 15.4; DB 8; Length 1664;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACCAACAAAGATTGCG 18
||||| ||||| ||||| |||||
DB 987 AGACTACAAAGATTGCG 971
AY692867 1950 bp DNA linear PLN 11-AUG-2004
Saccharomyces cerevisiae clone FLH158310.01X YIL155C gene, complete
cds.
AY692867 GI:51013184
AY692867 Yeast ORF Project.
Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
1 (bases 1 to 1950)
Marsischky,G., Rolfs,A., Richardson,A., Kane,M., Baqui,M.,
Taycher,E., Hu,Y., Vannberg,F., Weger,J., Kramer,J., Moreira,D.,
Kelley,F., Zhu,D., Raphael,J., Hogle,C., Jepson,D., Williamson,J.,
Canarso,A., Gonzaga,L., Vasconcelos,A.T., Simpson,A., Kolodner,R.,
Harlow,E. and Labaer,J.
Creation of the YFLEX clone resource: cloning of Saccharomyces
cerevisiae ORFs in the Gateway recombinational cloning system
Unpublished
2 (bases 1 to 1950)
Marsischky,G., Rolfs,A., Richardson,A., Kane,M., Baqui,M.,
Taycher,E., Hu,Y., Vannberg,F., Weger,J., Kramer,J., Moreira,D.,
Kelley,F., Zhu,D., Raphael,J., Hogle,C., Jepson,D., Williamson,J.,
Canarso,A., Gonzaga,L., Vasconcelos,A.T., Simpson,A., Kolodner,R.,
Harlow,E. and Labaer,J.
Direct Submission
Submitted (20-JUL-2004) Biological Chemistry and Molecular
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141, USA
This clone is part of a collection of Saccharomyces cerevisiae
full-length ORF clones generated by the Harvard Institute of
Proteomics. Each CDS has been cloned with its native stop codon.
The CDS has been directionally cloned using the Gateway cloning
system into the donor vectors pDONR 201 or pDONR 221. Additional
sequences in the clone: 'TCCAGTGACCAC' after the attL1 site and
before the 'ATG' (from Research Genetics primers used to amplify
the ORFs, including a Kozak consensus sequence).
'ATCCCGGATTCCTCATG' after the stop codon and before the attL2
site (from the Research Genetics primers used to amplify the ORFs).
Location/Qualifiers
1. .1950
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/clone="FLH158310.01X"
/lab_host="DH5alpha T1 resistant"
/db_xref="GI:51013185"
1. .1950
/product="YIL155C"
/codon_start=1
/protein_id="AAT92886.1"
/db_xref="GI:51013185"

FEATURES
source

mRNA
CDS

FEATURES
source

/translation="MFSVTRRAAGAAAAMATATGTYWTSQDRPLVNDPSYVQV
FPTAAPQVSRDLDLAKHQFDVLIIGGATGTCALDARGLNVALVEKGDFA
SGTSSKSRMHGGVRLKAFWFKQADLVIEALNERKILINTAPHLCTVPLLI
PIYTSQVQPIYMGCKFYDFPAGSOKLSYLLSKATVEKAPMLTIDNLKASLVVHD
GSFSDRLNATLAI TAVENGATVLYVEVQRLIKDPTSGKVI GAARDVETNELVRIN
AKCVNATGYPSSDAILQMDRNPGLPDSPLNDSKIKSTFNQIAVMDPKWVIPSIGVH

IVLPSFYCPKDMGLLDVRTSDGRVNFPLPMQGVLAGTTDIPLKQVFNPMPTBADIQ
DILKELQHYIFEPKVEDVLSAWGVPLVRDPTIPADGKKGSGATQGVVRSHPFLTS
DNGLITAGGKWTYRQMAESTVDKVVGVGFNLKPCHTPDILKAGAEWTONYVAL
LAQNYHLSKSMNVLQNGYRSGSIIIEFFKESMENKLPISLADKENNVIYSERNL
VNFDTFRYFTTIGELKYSQYCYKTPDLDFILRTRFAFLDADKALNAVAHATVKVMD
746
/note="compared to SGD sequence"
/replace="a"

variation

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1950;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||
Db 1795 AGAACAAAGATTCCG 1811
|||||

RESULT 39
AB091267 1997 bp mRNA linear PLN 04-MAR-2003
LOCUS
DEFINITION
Marchantia polymorpha mRNA for ftsZ1, complete cds.
ACCESSION
AB091267
VERSION
AB091267.1 GI:28804589
KEYWORDS
Marchantia polymorpha (liverwort)
ORGANISM
Marchantia polymorpha
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Marchantiophyta; Marchantiopsida; Marchantiidae; Marchantiales;
Marchantiaceae; Marchantiaceae; Marchantia.

REFERENCE
1
Araki,Y., Takio,S., Ono,K. and Takano,H.
Two types of plastid ftsZ genes in liverwort Marchantia polymorpha
Unpublished
2 (bases 1 to 1997)
Direct Submission
Takano,H., Araki,Y., Takio,S. and Ono,K.
Submitted (03-SEP-2002) Hiroyoshi Takano, kumamoto university,
Faculty of Science; Kurokumi, kumamoto-city, kumamoto 860-8555,
Japan (E-mail: takano@kumamoto-u.ac.jp).
URL: <http://www.sci.kumamoto-u.ac.jp/bio.iden/takano/english.html>,
Tel:81-96-342-3432, Fax:81-96-342-3432)

FEATURES
source
1..1997
/organism="Marchantia polymorpha"
/mol_type="mRNA"
/db_xref="taxon:3197"
1..1997
/gene="ftsZ1"
234..1574
/gene="ftsZ1"
/codon_start=1
/product="ftsZ1"
/protein_id="BAC57993.1"
/db_xref="GI:28804590"
/translation="MTVQVRPLTNSCHGSPVFASARPAICSGVGFARCCALPS
LHGKVSLLSPYLGSGTQHDCKRAEWQRAKSGATGTRIPASMI PMDSARIKVIQVG
GGNNAINMIGSLQGVFEWNTADQALLQSAATHRVQIGETLIRGLGTGNGPELG
EXAAEBSLEIAEAVSDADLVFITAGMGGTGGAAPVVARLAKEGGQITGVGVTFP
TTEGRRACQGLEAIBOLKQNDVTLIVINDELLDVVQEAFLIADLDVLRQGV
QGISITITPGVNVNDFADVGMNSGTAMLGVMSTGKRAEEAQAQATSAPLIER
STRATGVVNTIGKDKLTQEVNRVSQVVTGLADPAANIIFGAVVDEKVTGAVHTI
IATGFSQTQKTLIDPKVARQEQSPKGVDPVWRPAPVSRFPQGLSGKGLF"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1997;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||
Db 1795 AGAACAAAGATTCCG 1811
|||||

Db 1472 AGAACAAAGATTCCG 1488

RESULT 40
BD095617/c
LOCUS
DEFINITION
Process for biologically producing L-pipecolic acid.
ACCESSION
BD095617
VERSION
BD095617.1 GI:22641205
KEYWORDS
WO 0148216-A/8.
SOURCE
unidentified
ORGANISM
unidentified
unclassified.
REFERENCE
1 (bases 1 to 2186)
AUTHORS
Fujii,T., Aritoku,Y., Mukaiharu,M., Narita,T., Agematsu,H. and
Isshiki,K.
TITLE
Process for biologically producing L-pipecolic acid
JOURNAL
Patent: WO 0148216-A 8 05-JUL-2001;
MERCAN CORP,TADASHI FUJII,YASUHIRO ASITOKU,MANABU MUKAIHARA, TAKAO
NARITA, HITOSHI AGEMATSU,KUNIO ISSHIKI
OS Escherichia coli
PN WO 0148216-A/8
PD 05-JUL-2001
PF 22-DEC-2000 WO 2000JP009137
PR 28-DEC-1999 JP 99P 373389
PI TADASHI FUJII,YASUHIRO ASITOKU,MANABU MUKAIHARA,TAKAO NARITA,
HITOSHI AGEMATSU,KUNIO ISSHIKI
PC C12N15/54,C12N1/21,C12P17/12//C12N9/10,(C12N15/54,C12P17/12),
PC C12N1/21,C12P1/19),(C12N1/21,C12P1/15),(C12P17/12,C12P1/19),
PC (C12P17/12,C12P1/15),(C12N9/10,C12P1/20)
CC Process for biologically producing L-pipecolic acid FH Key
Location/Qualifiers
FT source 1..2186
FT /organism='Escherichia coli'.
Location/Qualifiers
1..2186
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 2186;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||
Db 1281 AGAACAAAGATTCCG 1265
|||||

RESULT 41
CQ716675 2845 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 2609 from Patent WO02068579.
ACCESSION
CQ716675
VERSION
CQ716675.1 GI:42277532
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 2609 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1..2845
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1997;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||
Db 1795 AGAACAAAGATTCCG 1811
|||||


```
FT source 1. .3178
FT Location/Qualifiers
FEATURES
  source
    1. .3178
      Location/Qualifiers
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 3178;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||||
Db 2154 AGAACACACAGATTGCG 2170

RESULT 46
LOCUS SCGUT2A
DEFINITION S. cerevisiae GUT2 gene.
ACCESSION X71660
VERSION X71660.1 GI:297115
KEYWORDS glycerol-3-phosphate dehydrogenase.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
  Bacteria: Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
  1 Ronnow, B. and Kielland-Brandt, M.C.
  GUT2, a gene for mitochondrial glycerol 3-phosphate dehydrogenase
  of Saccharomyces cerevisiae
  Yeast 9 (10), 1121-1130 (1993)
  MEDLINE 94078674
  PUBMED 8256521
REFERENCE
  2 (bases 1 to 3178)
  Roennow, B.
  Direct Submission
  Submitted (26-APR-1993) B. Roennow, Carlsberg Laboratory, Dept. of
  Physiology, Gamle Carlsberg Vej 10., DK-2500 Copenhagen Valby,
  DENMARK
FEATURES
  source
    1. .3178
      Location/Qualifiers
        /organism="Saccharomyces cerevisiae"
        /mol_type="genomic DNA"
        /strain="GRF88"
        /db_xref="taxon:4932"
        356..360
        /TATA_signal
        gene
          468..2312
          /gene="GUT2"
          468..2312
          /gene="GUT2"
          /EC_number="1.1.99.5"
          /codon_start=1
          /product="glycerol-3-phosphate dehydrogenase"
          /protein_id="CAA50652.1"
          /db_xref="GI:297116"
          /db_xref="GOA:P32191"
          /db_xref="Swiss-Prot:P32191"
          /translation="MTRATWNSPPPLHQRQVRRDLRLDKTHQFDVLIIGGATGT
          GCALDAATRLNALVEKDFASGTSKTKMTHGGVRYLEKAFWFEFSKAQDLVIEA
          LNERKLINTAPHLCTVLPILPIYSTVOVPVLYMGCKFYDFGGSNLKKSYLLSKS
          ATVEKAPMLTTNKLKSLVYHDSFNDSRLNATLAITGVENGATVLIYVEOKLIKDP
          TSKGVIGAEARDVETNELVRINAKCVVNATGPYSDALLOMDRNPGLPDSPLNDRSKI
          KSTFNISVMDPKWIPSGIHVILFSPSPKMDGLDVRTSDGRVMFPLPWGKVLK
          GTTDILPKQVFNPMPTFADIQDLPELQHYIEFFVKREDVLSAAGVPEPLVDRPTI
          PADGKKGAQGVVGRSHFTSDNGLITTAGKWTYRQMAETVDKVVEVGGFNKLK
          PCHTRDIKLAGEEWTONYVALLAQNHLSSKMSNLYQNYGTRRSIIICEFFKSMEN
          KFLSLADKNNVYSEENNLNFDTEYPTTIGELKYSQVYECRTPLDPLLRTRR
          FLFLDAKALNVAHTAVKMGDFENWSEKKRQWELEKTVNFQGRGV"
          transit_peptide
            468..527
            /gene="GUT2"
```


VERSION AX463544.1 GI:21886318
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 Duggan B.M., Yao M.G. and Griffin, J.A.
AUTHORS Secreted proteins.
TITLE Patent: WO 0248337-A 108 20-JUN-2002;
JOURNAL INCYTE GENOMICS INC (US)
FEATURES
source
1. .4717
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7503512CB1"
ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 4717;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACAAACAGATTCCG 18
||| ||||| ||||| |||||
Db 368 AGCACAAACAGATTCCG 384
Search completed: December 3, 2004, 03:06:23
Job time : 635.921 secs

This Page Blank (uspto)